

NON-HOMOGENEOUS $T1$ THEOREM FOR BI-PARAMETER SINGULAR INTEGRALS

TUOMAS HYTÖNEN AND HENRI MARTIKAINEN

ABSTRACT. We prove a non-homogeneous $T1$ theorem for certain bi-parameter singular integral operators. Moreover, we discuss the related non-homogeneous Journé's lemma and product BMO theory.

CONTENTS

1. Introduction	1
2. The main theorem	3
3. Strategy of the proof	9
4. Separated cubes	12
5. Adjacent cubes	19
6. Nested cubes	25
7. Mixed paraproducts	29
8. Necessity of $T1 \in \text{BMO}_{\text{prod}}(\mu)$	33
References	44

1. INTRODUCTION

This paper deals with singular integrals with two attributes: 'bi-parameter' and 'non-homogeneous'. Both of them have been investigated before but, as far as we know, only one at a time. Here, for the first time, we work in the simultaneous presence of both complications.

To be more precise, we study bi-parameter singular integrals T acting on some class of functions with product domain $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$. However, instead of the Lebesgue measure we equip \mathbb{R}^{n+m} with a product measure $\mu = \mu_n \times \mu_m$, where the measures μ_n and μ_m are only assumed to be upper doubling (a condition that is more general than the assumption $\mu_n(B(x, r)) \leq Cr^s$, $\mu_m(B(x, r)) \leq Cr^t$). We

2010 *Mathematics Subject Classification.* 42B20.

Key words and phrases. Bi-parameter singular integral, non-homogeneous analysis.

T.H. is supported by the European Union through the ERC Starting Grant Analytic-probabilistic methods for borderline singular integrals, and by the Academy of Finland, grants 130166 and 133264. H.M. is supported by the Emil Aaltonen Foundation, and by the Academy of Finland, grants 130166 and 133264. This work was started when H.M. was still at the University of Helsinki.

establish a $T1$ theorem, i.e. a boundedness criterion for singular integral operators, in this setting.

Such a result touches on quite a few topics. We now lay down some of the context. After the classical $T1$ and Tb type theory by David and Journé [4] and David, Journé and Semmes [5], the first $T1$ type theorem for product spaces was proved by Journé [14]. Journé formulated the assumptions in the language of vector-valued Calderón–Zygmund theory. Recently, Journé’s result was challenged by a new product-space $T1$ theorem by S. Pott and P. Villarroya [20], who introduced an alternative framework avoiding the vector-valued assumptions. Following a similar philosophy, the latter author proved a bi-parameter analog [16] of the representation theorem of the first author [9] – a particular consequence of which is yet another bi-parameter $T1$ theorem of a more dyadic flavour than [20].

All the classical multiparameter methods including Journé’s covering lemma [13] and the product BMO, Hardy space and multiparameter singular integral theory by Chang and Fefferman [1], [2], Fefferman [6] and Fefferman and Stein [7] are in the doubling situation. The same is true for the results of the previous paragraph, and for the other related modern developments like [3], [15], [21] and [22]. These deal with variations of Journé’s covering lemma, multiparameter paraproducts and dyadic versions of multiparameter function spaces. Since all the relevant results and definitions, both classical and modern, are in the doubling situation, we have to consider Journé’s lemma, (dyadic) product BMO, H^1 -BMO type duality results, singular integrals and multiparameter paraproducts all from the new perspective of general measures.

The following explains the specific need for these types of results. Even in the one-parameter theory of non-homogeneous singular integrals, one needs to replace the familiar BMO condition by

$$\int_I |f - f_I|^p d\mu \leq L\mu(\kappa I),$$

where a parameter $\kappa > 1$ specifies an expansion κI of the cube I ; this can have a much larger measure than I when μ is non-doubling. Accordingly, we need to define and study a new space $\text{BMO}_{\text{prod}}(\mu)$, which is used in the formulation of our main theorem. To show that the new condition $T1 \in \text{BMO}_{\text{prod}}(\mu)$ is necessary, we need a version of Journé’s covering lemma for general product measures. Finally, to handle some mixed paraproducts, we need a certain new H^1 -BMO type duality inequality.

Our basic proof strategy of the $T1$ theorem uses dyadic probabilistic techniques adapted to the bi-parameter situation. These were already used by the latter author in [16] but, again, only in the doubling case. The dyadic and probabilistic methods of non-homogeneous analysis were pioneered by Nazarov, Treil and Volberg (see e.g. the non-homogeneous Tb theorems [18] and [19]). However, these powerful tools are not widely used in multiparameter harmonic analysis, and some extra care is needed. In this regard the current paper is a continuation

of the recent developments (see e.g. [8], [9], [10], [11], [16], [17]) in the probabilistic methods.

The proof starts by expanding in a product Haar basis adapted to the general measures μ_n and μ_m . The summation in the bi-parameter case is rather massive and the non-homogeneous techniques are needed in various parts of the summation. All kinds of bad-boundary terms appear – also in some new mixed situations. Moreover, there is a wide variety of new non-homogeneous paraproducts, some of them related to the new space $\text{BMO}_{\text{prod}}(\mu)$.

Just like in [16] our operators are defined using the philosophy of Pott and Villarroya [20]. That is, we mostly avoid the language of vector-valued formulations used in the original work of Journé [14]. Of course, our kernel estimates are tied to the non-homogeneous measures.

2. THE MAIN THEOREM

We use this section to introduce the (somewhat lengthy) framework necessary for the formulation of our main theorem. We consider the following class of measures:

2.1. Definition (Upper doubling measures). Let $\lambda: \mathbb{R}^n \times (0, \infty) \rightarrow (0, \infty)$ be a function so that $r \mapsto \lambda(x, r)$ is non-decreasing and $\lambda(x, 2r) \leq C_\lambda \lambda(x, r)$ for all $x \in \mathbb{R}^n$ and $r > 0$. We say that a Borel measure μ in \mathbb{R}^n is upper doubling with the dominating function λ , if $\mu(B(x, r)) \leq \lambda(x, r)$ for all $x \in \mathbb{R}^n$ and $r > 0$. We set $d_\lambda = \log_2 C_\lambda$.

The property $\lambda(x, |x - y|) \sim \lambda(y, |x - y|)$ would be convenient. This is luckily something that can be arranged for free. In [12, Proposition 1.1] it is shown that $\Lambda(x, r) := \inf_{z \in \mathbb{R}^n} \lambda(z, r + |x - z|)$ satisfies that $r \mapsto \Lambda(x, r)$ is non-decreasing, $\Lambda(x, 2r) \leq C_\lambda \Lambda(x, r)$, $\mu(B(x, r)) \leq \Lambda(x, r)$, $\Lambda(x, r) \leq \lambda(x, r)$ and $\Lambda(x, r) \leq C_\lambda \Lambda(y, r)$ if $|x - y| \leq r$. Therefore, we may (and do) always assume that dominating functions λ satisfy the additional symmetry property $\lambda(x, r) \leq C \lambda(y, r)$ if $|x - y| \leq r$.

Henceforth, let $\mu = \mu_n \times \mu_m$, where μ_n and μ_m are upper doubling measures on \mathbb{R}^n and \mathbb{R}^m respectively. The corresponding dominating functions are denoted by λ_n and λ_m . We use ℓ^∞ metrics on \mathbb{R}^n and \mathbb{R}^m .

2.1. Bi-parameter Calderón–Zygmund operators. We study an a priori bounded linear operator $T: L^2(\mu) \rightarrow L^2(\mu)$. In addition to the usual adjoint T^* , we will be concerned with the *partial adjoint* T_1 defined by

$$\langle T_1(f_1 \otimes f_2), g_1 \otimes g_2 \rangle = \langle T(g_1 \otimes f_2), f_1 \otimes g_2 \rangle.$$

The philosophy of the following assumptions is that we impose symmetric conditions on all the four operators T, T^*, T_1 and T_1^* .

2.2. Assumption (Full kernel representation). If $f = f_1 \otimes f_2$ and $g = g_1 \otimes g_2$ with $f_1, g_1: \mathbb{R}^n \rightarrow \mathbb{C}$, $f_2, g_2: \mathbb{R}^m \rightarrow \mathbb{C}$, $\text{spt } f_1 \cap \text{spt } g_1 = \emptyset$ and $\text{spt } f_2 \cap \text{spt } g_2 = \emptyset$, we have

the kernel representation

$$\begin{aligned}\langle Tf, g \rangle &= \int_{\mathbb{R}^{n+m}} \int_{\mathbb{R}^{n+m}} K(x, y) f(y) g(x) d\mu(x) d\mu(y) \\ &= \iint_{\mathbb{R}^n \times \mathbb{R}^m} \iint_{\mathbb{R}^n \times \mathbb{R}^m} K(x_1, x_2; y_1, y_2) f(y_1, y_2) g(x_1, x_2) d\mu(x_1, x_2) d\mu(y_1, y_2),\end{aligned}$$

where the kernel K is a function

$$K: (\mathbb{R}^{n+m} \times \mathbb{R}^{n+m}) \setminus \{(x, y) \in \mathbb{R}^{n+m} \times \mathbb{R}^{n+m} : x_1 = y_1 \text{ or } x_2 = y_2\} \rightarrow \mathbb{C}.$$

Note that this implies kernel representations for T^* , T_1 and T_1^* . If we denote their kernels respectively by K^* , K_1 and K_1^* , then we immediately observe the formulae

$$\begin{aligned}K^*(x, y) &= K(y_1, y_2; x_1, x_2), \\ K_1(x, y) &= K(y_1, x_2; x_1, y_2), \\ K_1^*(x, y) &= K(x_1, y_2; y_1, x_2).\end{aligned}$$

2.3. Assumption (Full standard estimates). The kernel is assumed to satisfy the *size condition*

$$|K(x, y)| \leq C \frac{1}{\lambda_n(x_1, |x_1 - y_1|)} \frac{1}{\lambda_m(x_2, |x_2 - y_2|)},$$

the *Hölder condition*

$$\begin{aligned}&|K(x, y) - K(x, (y_1, y'_2)) - K(x, (y'_1, y_2)) + K(x, y')| \\ &\leq C \frac{|y_1 - y'_1|^\alpha}{|x_1 - y_1|^\alpha \lambda_n(x_1, |x_1 - y_1|)} \frac{|y_2 - y'_2|^\beta}{|x_2 - y_2|^\beta \lambda_m(x_2, |x_2 - y_2|)}\end{aligned}$$

whenever $|y_1 - y'_1| \leq |x_1 - y_1|/2$ and $|y_2 - y'_2| \leq |x_2 - y_2|/2$, and the *mixed Hölder and size condition*

$$|K(x, y) - K(x, (y'_1, y_2))| \leq C \frac{|y_1 - y'_1|^\alpha}{|x_1 - y_1|^\alpha \lambda_n(x_1, |x_1 - y_1|)} \frac{1}{\lambda_m(x_2, |x_2 - y_2|)}$$

whenever $|y_1 - y'_1| \leq |x_1 - y_1|/2$. The same conditions are imposed on K^* , K_1 and K_1^* as well.

We need to assume some Calderón–Zygmund structure on \mathbb{R}^n and \mathbb{R}^m separately.

2.4. Assumption (Partial kernel representation). If $f = f_1 \otimes f_2$ and $g = g_1 \otimes g_2$ with $\text{spt } f_1 \cap \text{spt } g_1 = \emptyset$, then we assume the kernel representation

$$\langle Tf, g \rangle = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} K_{f_2, g_2}(x_1, y_1) f_1(y_1) g_1(x_1) d\mu_n(x_1) d\mu_n(y_1).$$

2.5. Assumption (Partial boundedness \times standard estimates). The kernel

$$K_{f_2, g_2} : (\mathbb{R}^n \times \mathbb{R}^n) \setminus \{(x_1, y_1) \in \mathbb{R}^n \times \mathbb{R}^n : x_1 = y_1\}$$

is assumed to satisfy the size condition

$$|K_{f_2, g_2}(x_1, y_1)| \leq C(f_2, g_2) \frac{1}{\lambda_n(x_1, |x_1 - y_1|)}$$

and the Hölder conditions

$$|K_{f_2, g_2}(x_1, y_1) - K_{f_2, g_2}(x'_1, y_1)| \leq C(f_2, g_2) \frac{|x_1 - x'_1|^\alpha}{|x_1 - y_1|^\alpha \lambda_n(x_1, |x_1 - y_1|)}$$

whenever $|x_1 - x'_1| \leq |x_1 - y_1|/2$, and

$$|K_{f_2, g_2}(x_1, y_1) - K_{f_2, g_2}(x_1, y'_1)| \leq C(f_2, g_2) \frac{|y_1 - y'_1|^\alpha}{|x_1 - y_1|^\alpha \lambda_n(x_1, |x_1 - y_1|)}$$

whenever $|y_1 - y'_1| \leq |x_1 - y_1|/2$.

We assume that $C(f_2, g_2) \lesssim \|f_2\|_{L^2(\mu_n)} \|g_2\|_{L^2(\mu_m)}$.

We assume the analogous representation and properties with a kernel K_{f_1, g_1} in the case $\text{spt } f_2 \cap \text{spt } g_2 = \emptyset$.

2.6. Remark. As is clear, the final bound is not allowed to depend on $\|T\|_{L^2(\mu) \rightarrow L^2(\mu)}$. However, note that we do assume some type of L^2 boundedness separately on (\mathbb{R}^n, μ_n) and (\mathbb{R}^m, μ_m) (the bounds for $C(f_1, g_1)$ and $C(f_2, g_2)$). This is in slight contrast with some of the most recent works in the homogeneous bi-parameter setting [16]. There one may work with a bit weaker testing conditions for $C(\cdot, \cdot)$. However, this seems to be a standard assumption in the classical homogeneous works, and we also require it in our non-homogeneous setting.

2.7. Assumption (Full weak boundedness property). We assume the following weak boundedness property: for every cube $I \subset \mathbb{R}^n$ and $J \subset \mathbb{R}^m$ there holds that

$$|\langle T(\chi_I \otimes \chi_J), \chi_I \otimes \chi_J \rangle| \leq C\mu_n(5I)\mu_m(5J).$$

Note that this condition is invariant with respect to the replacement of T by either T^* , T_1 or T_1^* .

Moreover, we need some type of diagonal BMO assumptions. This amounts to replacing exactly one of the indicator functions above by a cancellative function supported on the same cube. More precisely:

2.8. Assumption (Partial weak boundedness \times BMO condition). Let $I \subset \mathbb{R}^n$ and $J \subset \mathbb{R}^m$ be cubes and a_I, a_J be functions such that $\int a_I d\mu_n = \int a_J d\mu_m = 0$, $\text{spt } a_I \subset I$ and $\text{spt } a_J \subset J$. We assume that

$$|\langle T(\chi_I \otimes \chi_J), a_I \otimes \chi_J \rangle| + |\langle T^*(\chi_I \otimes \chi_J), a_I \otimes \chi_J \rangle| \leq C\|a_I\|_{L^2(\mu_n)}\mu_n(5I)^{1/2}\mu_m(5J)$$

and

$$|\langle T(\chi_I \otimes \chi_J), \chi_I \otimes a_J \rangle| + |\langle T^*(\chi_I \otimes \chi_J), \chi_I \otimes a_J \rangle| \leq C\mu_n(5I)\|a_J\|_{L^2(\mu_m)}\mu_m(5J)^{1/2}.$$

Note that no new conditions arise by replacing T by T_1 above. The reason for calling this a BMO condition will become more clear in the course of the proof. Observe that both the full weak boundedness and the partial weak boundedness \times BMO condition are obviously necessary conditions for the boundedness of T on $L^2(\mu)$. In fact, they would both follow from the following stronger diagonal testing condition

$$\|\chi_I \otimes \chi_J T(\chi_I \otimes \chi_J)\|_{L^2(\mu)} + \|\chi_I \otimes \chi_J T^*(\chi_I \otimes \chi_J)\|_{L^2(\mu)} \leq C\mu_n(5I)^{1/2}\mu_m(5J)^{1/2}.$$

2.2. Dyadic grids and Haar functions. In order to formulate the actual “ $T_1 \in \text{BMO}$ ” conditions in the present setting, it is necessary to introduce some notation related to dyadic martingale difference decompositions and Haar functions.

Let \mathcal{D}_n^0 and \mathcal{D}_m^0 be the standard dyadic grids on \mathbb{R}^n and \mathbb{R}^m respectively.

2.2.1. Random dyadic grids and good/bad cubes. We let $w_n = (w_{n,i})_{i \in \mathbb{Z}}$, $w'_n = (w'_{n,i})_{i \in \mathbb{Z}}$, $w_m = (w_{m,j})_{j \in \mathbb{Z}}$ and $w'_m = (w'_{m,j})_{j \in \mathbb{Z}}$, where $w_{n,i}, w'_{n,i} \in \{0, 1\}^n$ and $w_{m,j}, w'_{m,j} \in \{0, 1\}^m$. In \mathbb{R}^n we define the new dyadic grid $\mathcal{D}_n = \{I + \sum_{i: 2^{-i} < \ell(I)} 2^{-i} w_{n,i} : I \in \mathcal{D}_n^0\} = \{I + w_n : I \in \mathcal{D}_n^0\}$, where we simply have defined $I + w_n := I + \sum_{i: 2^{-i} < \ell(I)} 2^{-i} w_{n,i}$. The grids \mathcal{D}'_n , \mathcal{D}_m and \mathcal{D}'_m are similarly defined. There is a natural product probability structure on $(\{0, 1\}^n)^\mathbb{Z}$ and $(\{0, 1\}^m)^\mathbb{Z}$. Even if $n = m$ and despite the notation we agree that we have four independent random dyadic grids.

A cube $I_1 \in \mathcal{D}_n$ is called bad if there exists $I_2 \in \mathcal{D}'_n$ so that $\ell(I_2) \geq 2^r \ell(I_1)$ and $d(I_1, \partial I_2) \leq 4\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}$. Here $\ell(I)$ denotes the side length of I and $\gamma_n = \alpha/(2d_{\lambda_n} + 2\alpha)$, where $\alpha > 0$ appears in the kernel estimates and, we recall, $d_{\lambda_n} = \log_2 C_{\lambda_n}$. The parameter $r > 0$ will be fixed later. Note also that here bad really means \mathcal{D}'_n -bad, but this is not usually spelled out. The badness in \mathbb{R}^m is defined similarly (it involves the same parameter r and $\gamma_m = \beta/(2d_{\lambda_m} + 2\beta)$).

It is easy to check that, almost surely with respect to the canonical probability on $(\{0, 1\}^n)^\mathbb{Z}$, a random dyadic system \mathcal{D}_n has the following “no quadrants” property: Whenever $I_k \in \mathcal{D}_n$, $k \in \mathbb{N}$, is a strictly increasing sequence (meaning that $I_k \subsetneq I_{k+1}$) of dyadic cubes, then this sequence exhausts all of \mathbb{R}^n (meaning that $\bigcup_{k \in \mathbb{N}} I_k = \mathbb{R}^n$). By throwing away a subset of the probability space with probability zero, we can and will assume that all our dyadic systems possess this additional “no quadrants” property. This convention is a matter of convenience, which has no effect on any of the probabilistic statements that we are going to make.

2.2.2. Haar functions on \mathbb{R}^n with general measures. The content of this subsection is from [10]. Let \mathcal{D}_n be one of the dyadic grids on \mathbb{R}^n as above and μ_n be any locally finite Borel measure on \mathbb{R}^n . We set

$$\begin{aligned} E_I f &= \langle f \rangle_I \chi_I, \\ \Delta_I f &= \sum_{I' \in \text{ch}(I)} [\langle f \rangle_{I'} - \langle f \rangle_I] \chi_{I'}, \end{aligned}$$

where $\text{ch}(I) = \{I_j : j = 1, \dots, 2^n\}$ is the collection of dyadic children of I , and $\langle f \rangle_I = \mu_n(I)^{-1} \int_I f d\mu_n$. For any $\ell \in \mathbb{Z}$ and $f \in L^2(\mu_n)$ we have the orthogonal decomposition

$$f = \sum_{\ell(I)=2^\ell} E_I f + \sum_{\ell(I) \leq 2^\ell} \Delta_I f.$$

We index the children I_j of I in such a way that

$$\mu_n(\widehat{I}_k) \geq [1 - (k-1)2^{-n}] \mu_n(I), \quad \widehat{I}_k = \bigcup_{u=k}^{2^n} I_u.$$

Hence $\mu_n(\widehat{I}_k) \sim \mu_n(I)$ for every k . We then continue to decompose

$$\begin{aligned} E_I f &= \langle f, h_{I,0} \rangle h_{I,0}, \\ \Delta_I f &= \sum_{\eta=1}^{2^n-1} \langle f, h_{I,\eta} \rangle h_{I,\eta}, \end{aligned}$$

where $h_{I,0} = \mu_n(I)^{-1/2} \chi_I$ and

$$h_{I,\eta} = \left(\frac{\mu_n(I_\eta) \mu_n(\widehat{I}_{\eta+1})}{\mu_n(\widehat{I}_\eta)} \right)^{1/2} \left[\frac{\chi_{I_\eta}}{\mu_n(I_\eta)} - \frac{\chi_{\widehat{I}_{\eta+1}}}{\mu_n(\widehat{I}_{\eta+1})} \right], \quad \eta = 1, \dots, 2^n - 1.$$

If $\mu_n(I_\eta) = 0$, we set $h_{I,\eta} = 0$.

We have the orthonormality $\langle h_{I_1,\eta_1}, h_{I_2,\eta_2} \rangle = \delta_{I_1}^{I_2} \delta_{\eta_1}^{\eta_2}$. Moreover, for $\eta \neq 0$ we have cancellation: $\int h_{I,\eta} d\mu_n = 0$. Finally, $h_{I,\eta}$ is supported on I and constant on its children. We have established that $\{h_{I,0} : \ell(I) = 2^\ell\} \cup \{h_{I,\eta} : \ell(I) \leq 2^\ell, \eta \in \{1, \dots, 2^n - 1\}\}$ is an orthonormal basis for $L^2(\mu_n)$ with the above mentioned properties.

Suppose now that $\mu = \mu_n \times \mu_m$, where μ_n and μ_m are locally finite Borel measures on \mathbb{R}^n and \mathbb{R}^m respectively. Assume also that we are given a dyadic grid \mathcal{D}_n on \mathbb{R}^n and a dyadic grid \mathcal{D}_m on \mathbb{R}^m . We denote the Haar functions on \mathbb{R}^n with respect to the measure μ_n by $h_{I,\eta}$, $I \in \mathcal{D}_n$, $\eta = 0, \dots, 2^n - 1$, and we denote the Haar functions on \mathbb{R}^m with respect to the measure μ_m by $u_{J,\kappa}$, $J \in \mathcal{D}_m$, $\kappa = 0, \dots, 2^m - 1$.

2.3. BMO conditions. We recall the definition of the BMO space relevant in the upper doubling setting from [19, 11].

2.9. Definition. We say that $f \in L^1_{\text{loc}}(\mu_n)$ belongs to $\text{BMO}_\kappa^p(\mu_n)$, if for any cube $I \subset \mathbb{R}^n$ there exists a constant f_I such that

$$\left(\int_I |f - f_I|^p d\mu_n \right)^{1/p} \leq L \mu_n(\kappa I)^{1/p},$$

where the constant L does not depend on I . The best constant L is denoted $\|f\|_{\text{BMO}_\kappa^p(\mu_n)}$.

The following lemma from [19, 11] motivates our definition of the BMO space for a product measure $\mu_n \times \mu_m$:

2.10. Lemma. For all $J \in \mathcal{D}'_n$,

$$\left(\sum_{\substack{I \in \mathcal{D}_n \text{ good} \\ I \subset J \\ \ell(I) \leq 2^{-r} \ell(J)}} |\langle f, h_I \rangle|^2 \right)^{1/2} \lesssim \mu_n(J)^{1/2} \|f\|_{\text{BMO}_\kappa^2(\mu_n)}.$$

2.11. Definition. We say that $b \in \text{BMO}_{\text{prod}}(\mu)$, if for every $\mathcal{D} = \mathcal{D}_n \times \mathcal{D}_m$ and $\mathcal{D}' = \mathcal{D}'_n \times \mathcal{D}'_m$ there holds that

$$\left(\sum_{\eta=1}^{2^n-1} \sum_{\kappa=1}^{2^m-1} \sum_{\substack{S \in \mathcal{D}' \\ S \subset \Omega}} \sum_{\substack{R=I \times J \in \mathcal{D} \text{ good}, R \subset S \\ \text{gen}(R) = \text{gen}(S) + (r, r)}} |\langle b, h_{I, \eta} \otimes u_{J, \kappa} \rangle|^2 \right)^{1/2} \leq L \mu(\Omega)^{1/2}$$

for all sets $\Omega \subset \mathbb{R}^{n+m}$ such that $\mu(\Omega) < \infty$ and such that for every $x \in \Omega$ there exists $S \in \mathcal{D}'$ so that $x \in S \subset \Omega$. The best constant L is denoted $\|b\|_{\text{BMO}_{\text{prod}}(\mu)}$. Here

$$\text{gen}(R) := (\text{gen}(I), \text{gen}(J)),$$

and $\text{gen}(I)$ is the usual generation of a dyadic cube (if $\ell(I) = 2^{-k}$ then $\text{gen}(I) = k$). Moreover, the fact that R is good means that I is \mathcal{D}'_n -good and J is \mathcal{D}'_m -good.

2.12. Remark. In the definition of $\text{BMO}_{\text{prod}}(\mu)$, we can equivalently restrict to considering sets Ω which, in addition to the stated property, are also bounded. Namely, if the defining inequality holds for this more restricted class of sets, and Ω is any set as in Definition 2.11, then we simply check the restricted product BMO condition for bounded sets of the form $\Omega \cap S_k$, where S_k is the union of the 2^{n+m} dyadic cubes in \mathcal{D}' of sidelength 2^k that lie closest to the origin. We conclude that the square sum over $S \subset \Omega \cap S_k$ is dominated by $L \mu(\Omega \cap S_k)^{1/2} \leq L \mu(\Omega)^{1/2}$, and taking the limit as $k \rightarrow \infty$, we find that the full sum over $S \subset \Omega$ is also dominated by $L \mu(\Omega)^{1/2}$, as required in Definition 2.11.

We can now formulate our main theorem:

2.13. Theorem. Let $\mu = \mu_n \times \mu_m$, where μ_n and μ_m are upper doubling measures on \mathbb{R}^n and \mathbb{R}^m respectively. Let $T: L^2(\mu) \rightarrow L^2(\mu)$ be a bi-parameter Calderón–Zygmund operator, which is a priori bounded. Assume that T satisfies all the Assumptions formulated in this section. Moreover, assume that

$$S1 \in \text{BMO}_{\text{prod}}(\mu) \quad \forall S \in \{T, T^*, T_1, T_1^*\}.$$

Then there holds that $\|T\| \lesssim 1$, a bound depending only on the Assumptions and the $\text{BMO}_{\text{prod}}(\mu)$ norms of the four $S1$, but independent of the a priori bound.

2.14. Remark. Note that at least for a large subclass of operators (see Section 8) our conditions, except for the conditions $T_1(1), T_1^*(1) \in \text{BMO}_{\text{prod}}(\mu)$, are necessary. Indeed, T_1 does not have to be bounded even if T is. This deficit is shared by all known bi-parameter $T1$ theorems, already in the homogeneous setting.

3. STRATEGY OF THE PROOF

3.1. Initial reductions. The following reduction is quite standard, but we do it carefully, since in the bi-parameter case the details are not written anywhere else. We do not need an exact averaging equality as in some of the latest non-homogeneous work – the by now classical trick of Nazarov, Treil and Volberg suffices here.

Let us fix $f, g \in C(\mathbb{R}^{n+m})$ with compact support so that $0.7\|T\| \leq |\langle Tf, g \rangle|$ (here $\|T\| = \|T\|_{L^2(\mu) \rightarrow L^2(\mu)} < \infty$) and $\|f\|_{L^2(\mu)} = \|g\|_{L^2(\mu)} = 1$. We fix the parameter ℓ so that $\text{spt } f, \text{spt } g \subset [-2^\ell, 2^\ell]^{n+m}$.

We define

$$\begin{aligned} E_k f &= \sum_{\eta_1} \sum_{\substack{I_1 \in \mathcal{D}_n \\ 2^{-k} < \ell(I_1) \leq 2^\ell}} h_{I_1, \eta_1} \otimes \langle f, h_{I_1, \eta_1} \rangle_1; \\ E'_k g &= \sum_{\eta_2} \sum_{\substack{I_2 \in \mathcal{D}'_n \\ 2^{-k} < \ell(I_2) \leq 2^\ell}} h_{I_2, \eta_2} \otimes \langle g, h_{I_2, \eta_2} \rangle_1; \\ \tilde{E}_k f &= \sum_{\kappa_1} \sum_{\substack{J_1 \in \mathcal{D}_m \\ 2^{-k} < \ell(J_1) \leq 2^\ell}} \langle f, u_{J_1, \kappa_1} \rangle_2 \otimes u_{J_1, \kappa_1}; \\ \tilde{E}'_k g &= \sum_{\kappa_2} \sum_{\substack{J_2 \in \mathcal{D}'_m \\ 2^{-k} < \ell(J_2) \leq 2^\ell}} \langle g, u_{J_2, \kappa_2} \rangle_2 \otimes u_{J_2, \kappa_2}. \end{aligned}$$

Here $\langle f, h_{I_1, \eta_1} \rangle_1(y_2) = \langle f(\cdot, y_2), h_{I_1, \eta_1} \rangle = \int f(y_1, y_2) h_{I_1, \eta_1}(y_1) d\mu_n(y_1)$.

Note that these are linear operators, $f = \lim_{k \rightarrow \infty} E_k f = \lim_{k \rightarrow \infty} \tilde{E}_k f$ on $L^2(\mu)$, $\|E_k f\|_{L^2(\mu)}, \|\tilde{E}_k f\|_{L^2(\mu)} \leq \|f\|_{L^2(\mu)} = 1$ and $E_k \tilde{E}_k f = \tilde{E}_k E_k f$. We write

$$\begin{aligned} \langle Tf, g \rangle &= \langle T(f - E_k f), g \rangle + \langle T(E_k f), g - E'_k g \rangle \\ &\quad + \langle T(E_k f - \tilde{E}_k E_k f), E'_k g \rangle + \langle T(\tilde{E}_k E_k f), E'_k g - \tilde{E}'_k E'_k g \rangle \\ &\quad + \langle T(\tilde{E}_k E_k f), \tilde{E}'_k E'_k g \rangle, \end{aligned}$$

or more briefly, $\langle Tf, g \rangle = \langle T(\tilde{E}_k E_k f), \tilde{E}'_k E'_k g \rangle + \epsilon_k(\omega)$, $\omega = (\omega_n, \omega'_n, \omega_m, \omega'_m)$. It is easy to see using the above listed properties that $\lim_{k \rightarrow \infty} \epsilon_k(\omega) = 0$ and $|\epsilon_k(\omega)| \leq 2\|T\|$. Using dominated convergence we deduce that

$$\langle Tf, g \rangle = \lim_{k \rightarrow \infty} E \langle T(\tilde{E}_k E_k f), \tilde{E}'_k E'_k g \rangle,$$

where the expectation $E = E_\omega$.

Note that

$$\tilde{E}_k E_k f = \sum_{\eta_1, \kappa_1} \sum_{\substack{I_1 \in \mathcal{D}_n \\ 2^{-k} < \ell(I_1) \leq 2^\ell}} \sum_{\substack{J_1 \in \mathcal{D}_m \\ 2^{-k} < \ell(J_1) \leq 2^\ell}} \langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}.$$

We write

$$\begin{aligned} \sum_{\substack{I_1 \in \mathcal{D}_n \\ 2^{-k} < \ell(I_1) \leq 2^\ell}} \sum_{\substack{J_1 \in \mathcal{D}_m \\ 2^{-k} < \ell(J_1) \leq 2^\ell}} &= \sum_{\substack{I_1 \in \mathcal{D}_{n, \text{good}} \\ 2^{-k} < \ell(I_1) \leq 2^\ell}} \sum_{\substack{J_1 \in \mathcal{D}_{m, \text{good}} \\ 2^{-k} < \ell(J_1) \leq 2^\ell}} \\ &+ \sum_{\substack{I_1 \in \mathcal{D}_{n, \text{good}} \\ 2^{-k} < \ell(I_1) \leq 2^\ell}} \sum_{\substack{J_1 \in \mathcal{D}_{m, \text{bad}} \\ 2^{-k} < \ell(J_1) \leq 2^\ell}} + \sum_{\substack{I_1 \in \mathcal{D}_{n, \text{bad}} \\ 2^{-k} < \ell(I_1) \leq 2^\ell}} \sum_{\substack{J_1 \in \mathcal{D}_m \\ 2^{-k} < \ell(J_1) \leq 2^\ell}}. \end{aligned}$$

Denoting $f_k = \tilde{E}_k E_k f$, this gives us the decomposition

$$f_k = f_{k, \text{good}} + f_{k, \text{bad}}.$$

Let also $g_k = \tilde{E}'_k E'_k g = g_{k, \text{good}} + g_{k, \text{bad}}$.

We now write

$$\langle T f_k, g_k \rangle = \langle T f_{k, \text{good}}, g_{k, \text{good}} \rangle + \langle T f_{k, \text{good}}, g_{k, \text{bad}} \rangle + \langle T f_{k, \text{bad}}, g_k \rangle.$$

This gives us that

$$\begin{aligned} |\langle T f, g \rangle| &= \lim_{k \rightarrow \infty} |E \langle T f_k, g_k \rangle| \\ &\leq \lim_{k \rightarrow \infty} (|E \langle T f_{k, \text{good}}, g_{k, \text{good}} \rangle| + \|T\| \|E\| \|g_{k, \text{bad}}\|_{L^2(\mu)} + \|T\| \|E\| \|f_{k, \text{bad}}\|_{L^2(\mu)}). \end{aligned}$$

Estimating

$$\begin{aligned} \|f_{k, \text{bad}}\|_{L^2(\mu)} &\leq \left(\sum_{\eta_1, \kappa_1} \sum_{\substack{I_1 \in \mathcal{D}_n \\ \ell(I_1) \leq 2^\ell}} \sum_{\substack{J_1 \in \mathcal{D}_m \\ \ell(J_1) \leq 2^\ell}} \chi_{\text{bad}}(J_1) |\langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle|^2 \right)^{1/2} \\ &+ \left(\sum_{\eta_1, \kappa_1} \sum_{\substack{I_1 \in \mathcal{D}_n \\ \ell(I_1) \leq 2^\ell}} \sum_{\substack{J_1 \in \mathcal{D}_m \\ \ell(J_1) \leq 2^\ell}} \chi_{\text{bad}}(I_1) |\langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle|^2 \right)^{1/2} \end{aligned}$$

gives us the bound

$$\begin{aligned} &E \|f_{k, \text{bad}}\|_{L^2(\mu)} \\ &\leq E_{(\omega_n, \omega'_n, \omega_m)} \left(\sum_{\eta_1, \kappa_1} \sum_{\substack{I_1 \in \mathcal{D}_n \\ \ell(I_1) \leq 2^\ell}} \sum_{\substack{J_1 \in \mathcal{D}_m \\ \ell(J_1) \leq 2^\ell}} \mathbb{P}_{\omega'_m}(J_1 \text{ bad}) |\langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle|^2 \right)^{1/2} \\ &+ E_{(\omega_n, \omega_m, \omega'_m)} \left(\sum_{\eta_1, \kappa_1} \sum_{\substack{I_1 \in \mathcal{D}_n \\ \ell(I_1) \leq 2^\ell}} \sum_{\substack{J_1 \in \mathcal{D}_m \\ \ell(J_1) \leq 2^\ell}} \mathbb{P}_{\omega'_n}(I_1 \text{ bad}) |\langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle|^2 \right)^{1/2} \leq c(r), \end{aligned}$$

where $c(r) \rightarrow 0$, when $r \rightarrow \infty$. Fixing r to be sufficiently large we have shown that

$$(3.1) \quad 0.7 \|T\| \leq |\langle T f, g \rangle| \leq \lim_{k \rightarrow \infty} |E \langle T f_{k, \text{good}}, g_{k, \text{good}} \rangle| + 0.1 \|T\|.$$

3.2. Outline of the core of the proof. The core of the proof, which will span most of the rest of the paper, consists of showing that

$$\begin{aligned}
 & |E\langle T f_{k,\text{good}}, g_{k,\text{good}} \rangle| \\
 &= \left| E \sum_{\substack{\eta_1, \eta_2 \\ \kappa_1, \kappa_2}} \sum_{\substack{I_1, I_2 \text{ good} \\ 2^{-k} < \ell(I_1), \ell(I_2) \leq 2^\ell}} \sum_{\substack{J_1, J_2 \text{ good} \\ 2^{-k} < \ell(J_1), \ell(J_2) \leq 2^\ell}} \langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle \right. \\
 &\quad \left. \times \langle g, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle \right| \\
 &\leq (C + 0.1\|T\|)\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)} = C + 0.1\|T\|,
 \end{aligned} \tag{3.2}$$

and hence, combining with (3.1),

$$\begin{aligned}
 0.6\|T\| &\leq C + 0.1\|T\|, & \text{which implies} \\
 \|T\| &\leq 2C,
 \end{aligned}$$

completing the proof of Theorem 2.13.

For the proof of (3.2), we fix k (recall that ℓ is already fixed). We will bound $|E\langle T f_{k,\text{good}}, g_{k,\text{good}} \rangle|$ uniformly on these quantities. We suppress the restrictions

$$2^{-k} < \ell(I_1), \ell(I_2), \ell(J_1), \ell(J_2) \leq 2^\ell$$

from the notation. The goodness of the cubes is essential and so is the fact that we have the averaging operator E in front (this is used in the proof to control multiple bad boundary regions). However, sometimes (for the collapse of paraproducts) the fact that all the cubes are good is a problem. We deal with this during the proof by sometimes explicitly replacing f by f_{good} noting that $\langle f_{\text{good}}, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle = \langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle$ if I_1 and J_1 are good, and $\langle f_{\text{good}}, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle = 0$ otherwise.

We will focus on the part of the summation, where $\ell(I_1) \leq \ell(I_2)$ and $\ell(J_1) \leq \ell(J_2)$. One exception of this is made during the proof to handle a certain mixed full paraproduct appearing in the summation $\ell(I_1) \leq \ell(I_2), \ell(J_1) > \ell(J_2)$. It has an essential difference to the full paraproduct appearing in our main summation, so it does need separate attention.

In any case, we perform the splitting

$$\sum_{\ell(I_1) \leq \ell(I_2)} = \sum_{\substack{\ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) > 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}}} + \sum_{\substack{\ell(I_1) < 2^{-r} \ell(I_2) \\ d(I_1, I_2) \leq 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}}} + \sum_{\substack{2^{-r} \ell(I_2) \leq \ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) \leq 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}}}.$$

These three parts are called separated, nested and adjacent respectively. The term nested makes sense, since the summing conditions (recalling that I_1 is good) actually imply that there is a child $I_{2,1} \in \text{ch}(I_2)$ so that $d(I_1, I_{2,1}^c) > 4\ell(I_1)^{\gamma_n} \ell(I_{2,1})^{1-\gamma_n}$.

A similar splitting in the summation $\ell(J_1) \leq \ell(J_2)$ is also performed. This splits the whole summation into nine parts. We explicitly deal with the following cases (the remaining three being symmetric to one of these):

(J_1, J_2) (I_1, I_2)	separated	adjacent	nested
separated	*	*	*
adjacent		*	*
nested			*

The cases where the first pair (I_1, I_2) is separated are treated in Section 4, and the cases where it is adjacent in Section 5. The remaining cases where both pairs are nested are handled in Sections 6 and 7. A combination of these gives the core estimate (3.2) and completes the proof of Theorem 2.13. In the final Section 8, we investigate the necessity of the condition that $T1 \in \text{BMO}_{\text{prod}}(\mu)$.

3.3. Remark (Vinogradov notation and implicit constants). We will use the notation $f \lesssim g$ synonymously with $f \leq Cg$ for some constant C . We also use $f \sim g$ if $f \lesssim g \lesssim f$. The dependence on the various parameters should be somewhat clear, but basically C may depend on the various constants involved in the assumptions.

4. SEPARATED CUBES

In this section we deal with the part of the summation, where at least the pair of cubes (I_1, I_2) is separated. This further splits into subcases according to the nature of the other pair (J_1, J_2) .

4.1. Separated/separated.

4.1. Lemma. *Let $I_1 \in \mathcal{D}_n$, $I_2 \in \mathcal{D}'_n$, $J_1 \in \mathcal{D}_m$ and $J_2 \in \mathcal{D}'_m$ be such that $\ell(I_1) \leq \ell(I_2)$, $\ell(J_1) \leq \ell(J_2)$, $d(I_1, I_2) > 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}$ and $d(J_1, J_2) > 2\ell(J_1)^{\gamma_m} \ell(J_2)^{1-\gamma_m}$. Let $f_{I_1}, f_{J_1}, g_{I_2}$ and g_{J_2} be functions supported on I_1, J_1, I_2 and J_2 respectively, and assume that $\|f_{I_1}\|_{L^2(\mu_n)} = \|f_{J_1}\|_{L^2(\mu_m)} = \|g_{I_2}\|_{L^2(\mu_n)} = \|g_{J_2}\|_{L^2(\mu_m)} = 1$ and $\int f_{I_1} d\mu_n = \int f_{J_1} d\mu_m = 0$. Then there holds that*

$$|\langle T(f_{I_1} \otimes f_{J_1}), g_{I_2} \otimes g_{J_2} \rangle| \lesssim A_{I_1 I_2}^{\text{sep}} A_{J_1 J_2}^{\text{sep}},$$

where

$$A_{I_1 I_2}^{\text{sep}} = \frac{\ell(I_1)^{\alpha/2} \ell(I_2)^{\alpha/2}}{D(I_1, I_2)^\alpha \sup_{z \in I_1} \lambda_n(z, D(I_1, I_2))} \mu_n(I_1)^{1/2} \mu_n(I_2)^{1/2}$$

and

$$A_{J_1 J_2}^{\text{sep}} = \frac{\ell(J_1)^{\beta/2} \ell(J_2)^{\beta/2}}{D(J_1, J_2)^\beta \sup_{w \in J_1} \lambda_m(w, D(J_1, J_2))} \mu_m(J_1)^{1/2} \mu_m(J_2)^{1/2}.$$

Here $D(I_1, I_2) = \ell(I_1) + \ell(I_2) + d(I_1, I_2)$.

Proof. We write $\langle T(f_{I_1} \otimes f_{J_1}), g_{I_2} \otimes g_{J_2} \rangle$ as the integral

$$\int_{I_1} \int_{J_1} \int_{I_2} \int_{J_2} K(x, y) f_{I_1}(y_1) f_{J_1}(y_2) g_{I_2}(x_1) g_{J_2}(x_2) d\mu_m(x_2) d\mu_n(x_1) d\mu_m(y_2) d\mu_n(y_1)$$

and then replace $K(x, y)$ by

$$K(x, y) - K(x, (y_1, w)) - K(x, (z, y_2)) + K(x, (z, w)),$$

where $z \in I_1$ and $w \in J_1$ are arbitrary. The replacement can be done inside the integral, since $\int f_{I_1} d\mu_n = \int f_{J_1} d\mu_m = 0$.

Next, one notes that

$$|y_1 - z| \leq \ell(I_1) \leq \ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n} \leq d(I_1, I_2)/2 \leq |x_1 - z|/2,$$

and that similarly $|y_2 - w| \leq |x_2 - w|/2$. Using the Hölder condition of K this yields

$$\begin{aligned} & |K(x, y) - K(x, (y_1, w)) - K(x, (z, y_2)) + K(x, (z, w))| \\ & \lesssim \frac{|y_1 - z|^\alpha}{|x_1 - z|^\alpha \lambda_n(z, |x_1 - z|)} \frac{|y_2 - w|^\beta}{|x_2 - w|^\beta \lambda_m(w, |x_2 - w|)} \\ & \leq \frac{\ell(I_1)^\alpha}{d(I_1, I_2)^\alpha \lambda_n(z, d(I_1, I_2))} \frac{\ell(J_1)^\beta}{d(J_1, J_2)^\beta \lambda_m(w, d(J_1, J_2))}. \end{aligned}$$

Splitting into cases $d(I_1, I_2) \geq \ell(I_2)$ and $d(I_1, I_2) < \ell(I_2)$ one can show precisely like in the proof of [11, Lemma 6.2] that

$$\frac{\ell(I_1)^\alpha}{d(I_1, I_2)^\alpha \lambda_n(z, d(I_1, I_2))} \lesssim \frac{\ell(I_1)^{\alpha/2} \ell(I_2)^{\alpha/2}}{D(I_1, I_2)^\alpha \lambda_n(z, D(I_1, I_2))}.$$

The result of the lemma readily follows. \square

We recall [11, Proposition 6.3]:

4.2. Proposition. *There holds*

$$\sum_{\ell(I_1) \leq \ell(I_2)} A_{I_1 I_2}^{\text{sep}} x_{I_1} y_{I_2} \lesssim \left(\sum_{I_1} x_{I_1}^2 \right)^{1/2} \left(\sum_{I_2} y_{I_2}^2 \right)^{1/2}$$

for $x_{I_1}, y_{I_2} \geq 0$. In particular, there holds that

$$\left(\sum_{I_2} \left[\sum_{I_1: \ell(I_1) \leq \ell(I_2)} A_{I_1 I_2}^{\text{sep}} x_{I_1} \right]^2 \right)^{1/2} \lesssim \left(\sum_{I_1} x_{I_1}^2 \right)^{1/2}.$$

We are to bound

$$\begin{aligned} & \sum_{\substack{\eta_1, \eta_2 \\ \kappa_1, \kappa_2}} \sum_{\substack{\ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) > 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}}} \sum_{\substack{\ell(J_1) \leq \ell(J_2) \\ d(J_1, J_2) > 2\ell(J_1)^{\gamma_m} \ell(J_2)^{1-\gamma_m}}} |\langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle| \\ & \quad \times |\langle g, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle| |\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle|. \end{aligned}$$

In general, we need to split

$$\sum_{\substack{\eta_1, \eta_2 \\ \kappa_1, \kappa_2}} = \sum_{\substack{\eta_1=0 \\ \kappa_1=0}} + \sum_{\substack{\eta_1=0 \\ \kappa_1 \neq 0}} + \sum_{\substack{\eta_1 \neq 0 \\ \kappa_1=0}} + \sum_{\substack{\eta_1 \neq 0 \\ \kappa_1 \neq 0}}.$$

However, in this separated/separated sum only the fourth summation appears. Indeed, say $\eta_1 = 0$. Then $2^\ell = \ell(I_1) \leq \ell(I_2) \leq 2^\ell$ i.e. $\ell(I_1) = \ell(I_2) = 2^\ell$. If the

cubes I_1 and I_2 actually contribute to the sum, one must have $I_1 \cap [-2^\ell, 2^\ell]^n \neq \emptyset$ and $I_2 \cap [-2^\ell, 2^\ell]^n \neq \emptyset$. But this means that $d(I_1, I_2) \leq 2^{\ell+1} = 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}$.

The main term, where $\eta_1 \neq 0$ and $\kappa_1 \neq 0$, is the only one that remains and can now be handled. In that sum Lemma 4.1 gives

$$|\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle| \lesssim A_{I_1 I_2}^{\text{sep}} A_{J_1 J_2}^{\text{sep}}.$$

With fixed $\eta_1 \neq 0$, $\kappa_1 \neq 0$, η_2 and κ_2 we get from Proposition 4.2 that

$$\begin{aligned} & \sum_{\ell(I_1) \leq \ell(I_2)} \sum_{\ell(J_1) \leq \ell(J_2)} A_{I_1 I_2}^{\text{sep}} A_{J_1 J_2}^{\text{sep}} |\langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle| |\langle g, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle| \\ & \lesssim \sum_{\ell(I_1) \leq \ell(I_2)} A_{I_1 I_2}^{\text{sep}} \left(\sum_{J_1} |\langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle|^2 \right)^{1/2} \left(\sum_{J_2} |\langle g, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle|^2 \right)^{1/2} \\ & \lesssim \left(\sum_{I_1} \sum_{J_1} |\langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle|^2 \right)^{1/2} \left(\sum_{I_2} \sum_{J_2} |\langle g, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle|^2 \right)^{1/2} \\ & \lesssim \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)}. \end{aligned}$$

4.2. Separated/nested. We begin by remarking that in this sum we automatically have $\eta_1 \neq 0$ and $\kappa_1 \neq 0$. We write $\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle$ as the sum

$$\begin{aligned} & [\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle - \langle u_{J_2, \kappa_2} \rangle_{J_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes 1 \rangle] \\ & + \langle u_{J_2, \kappa_2} \rangle_{J_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes 1 \rangle. \end{aligned}$$

We start by dealing with the sum having the first term as the matrix element. Let $J_{2,1} \in \text{ch}(J_2)$ be such that $J_1 \subset J_{2,1}$. Note that $\langle u_{J_2, \kappa_2} \rangle_{J_1} = \langle u_{J_2, \kappa_2} \rangle_{J_{2,1}}$. Therefore, we have that

$$\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle - \langle u_{J_2, \kappa_2} \rangle_{J_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes 1 \rangle$$

equals

$$\begin{aligned} & - \langle u_{J_2, \kappa_2} \rangle_{J_{2,1}} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes \chi_{J_{2,1}^c} \rangle \\ & + \sum_{\substack{J'_2 \in \text{ch}(J_2) \\ J'_2 \subset J_2 \setminus J_{2,1}}} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \chi_{J'_2} \rangle. \end{aligned}$$

As J_1 is good, $J_1 \subset J_{2,1}$ and $\ell(J_1) \leq 2^{-r} \ell(J_{2,1})$, we have

$$d(J_1, J_{2,1}^c) \geq 4\ell(J_1)^{\gamma_m} \ell(J_{2,1})^{1-\gamma_m} > 2\ell(J_1)^{\gamma_m} \ell(J_2)^{1-\gamma_m}.$$

4.3. Lemma. *There holds*

$$|\langle u_{J_2, \kappa_2} \rangle_{J_{2,1}} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes \chi_{J_{2,1}^c} \rangle| \lesssim A_{I_1 I_2}^{\text{sep}} A_{J_1 J_2}^{\text{in}},$$

where

$$A_{J_1 J_2}^{\text{in}} = \left(\frac{\ell(J_1)}{\ell(J_2)} \right)^{\beta/2} \left(\frac{\mu_m(J_1)}{\mu_m(J_{2,1})} \right)^{1/2}.$$

Proof. We write $\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes \chi_{J_{2,1}^c} \rangle$ as the integral

$$\int_{I_1} \int_{J_1} \int_{I_2} \int_{J_{2,1}^c} K(x, y) h_{I_1, \eta_1}(y_1) u_{J_1, \kappa_1}(y_2) h_{I_2, \eta_2}(x_1) d\mu_m(x_2) d\mu_n(x_1) d\mu_m(y_2) d\mu_n(y_1),$$

and then, using the fact that $\eta_1 \neq 0$ and $\kappa_1 \neq 0$, replace $K(x, y)$ with

$$K(x, y) - K(x, (y_1, w)) - K(x, (z, y_2)) + K(x, (z, w)).$$

Here $z \in I_1$ and $w \in J_1$ are arbitrary. Since $d(I_1, I_2) > 2\ell(I_1)^{\gamma_m} \ell(I_2)^{1-\gamma_m}$ and $d(J_1, J_{2,1}^c) > 2\ell(J_1)^{\gamma_m} \ell(J_2)^{1-\gamma_m}$, we have $|y_1 - z| \leq |x_1 - z|/2$ and $|y_2 - w| \leq |x_2 - w|/2$. Now, we have

$$\begin{aligned} |\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes \chi_{J_{2,1}^c} \rangle| &\lesssim A_{I_1 I_2}^{\text{sep}} \ell(J_1)^\beta \mu_m(J_1)^{1/2} \\ &\quad \times \int_{\mathbb{R}^m \setminus B(w, d(J_1, J_{2,1}^c))} \frac{|x_2 - w|^{-\beta}}{\lambda_m(w, |x_2 - w|)} d\mu_m(x_2), \end{aligned}$$

where

$$\int_{\mathbb{R}^m \setminus B(w, d(J_1, J_{2,1}^c))} \frac{|x_2 - w|^{-\beta}}{\lambda_m(w, |x_2 - w|)} d\mu_m(x_2) \lesssim d(J_1, J_{2,1}^c)^{-\beta} \lesssim \ell(J_1)^{-\beta/2} \ell(J_2)^{-\beta/2}.$$

Noting that $|\langle u_{J_2, \kappa_2} \rangle_{J_{2,1}}| \leq \mu_m(J_{2,1})^{-1/2}$ finishes the proof. \square

4.4. Lemma. *If $J'_2 \in \text{ch}(J_2)$ and $J'_2 \subset J_2 \setminus J_{2,1}$, then*

$$|\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \chi_{J'_2} \rangle| \lesssim A_{I_1 I_2}^{\text{sep}} A_{J_1 J_2}^{\text{in}}.$$

Proof. Since $d(J_1, J'_2) \geq d(J_1, J_{2,1}^c) > 2\ell(J_1)^{\gamma_m} \ell(J'_2)^{1-\gamma_m}$, Lemma 4.1 gives that

$$|\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \chi_{J'_2} \rangle| \lesssim A_{I_1 I_2}^{\text{sep}} \left(\frac{\ell(J_1)}{\ell(J_2)} \right)^{\beta/2} \mu_m(J_1)^{1/2} \frac{\mu_m(J'_2)^{1/2}}{\lambda_m(w, \ell(J'_2))}$$

with any $w \in J_1 \subset J_{2,1}$. It is easy to see that

$$\frac{\mu_m(J'_2)^{1/2}}{\lambda_m(w, \ell(J'_2))} \lesssim \mu_m(J_{2,1})^{-1/2},$$

which ends the proof of the lemma. \square

We recall [19, Lemma 7.4]:

4.5. Lemma. *There holds*

$$\sum_{\substack{J_1, J_2 \\ J_1 \subset J_{2,1} \in \text{ch}(J_2) \\ \ell(J_1) < 2^{-r} \ell(J_2)}} A_{J_1 J_2}^{\text{in}} x_{J_1} y_{J_2} \lesssim \left(\sum_{J_1} x_{J_1}^2 \right)^{1/2} \left(\sum_{J_2} y_{J_2}^2 \right)^{1/2}$$

for $x_{J_1}, y_{J_2} \geq 0$. In particular, there holds that

$$\left(\sum_{J_2} \left[\sum_{\substack{J_1 \subset J_2, 1 \in \text{ch}(J_2) \\ \ell(J_1) < 2^{-r} \ell(J_2)}} A_{J_1 J_2}^{\text{in}} x_{J_1} \right]^2 \right)^{1/2} \lesssim \left(\sum_{J_1} x_{J_1}^2 \right)^{1/2}.$$

Combining this lemma with Proposition 4.2 and the above estimates gives that the summation

$$\sum_{\substack{\ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) > 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}}} \sum_{\substack{\ell(J_1) < 2^{-r} \ell(J_2) \\ d(J_1, J_2) \leq 2\ell(J_1)^{\gamma_m} \ell(J_2)^{1-\gamma_m}}}$$

is dominated by $\|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)}$, if we are using the modified matrix element

$$\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle - \langle u_{J_2, \kappa_2} \rangle_{J_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes 1 \rangle.$$

What is left to do is to consider the same summation but with the matrix element

$$\langle u_{J_2, \kappa_2} \rangle_{J_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes 1 \rangle.$$

Recall that all the appearing cubes are good. In particular, J_2 is good and this is actually slightly problematic. However, for good I_2 and J_2 we can write $\langle g, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle = \langle g_{\text{good}}, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle$, and after this we may add all the bad J_2 to the summation, since $\langle g_{\text{good}}, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle = 0$ if J_2 is bad. Moreover, in the non-trivial case $\langle u_{J_2, \kappa_2} \rangle_{J_1} \neq 0$, we have $J_1 \cap J_2 \neq \emptyset$ so that we may remove the summing condition $d(J_1, J_2) \leq 2\ell(J_1)^{\gamma_m} \ell(J_2)^{1-\gamma_m}$.

In what follows we explicitly write the fact that I_1, I_2 and J_1 are good, since we need to highlight the fact that J_2 is not. For the J_2 summation we also write out the hidden summing condition $\ell(J_2) \leq 2^\ell$. For fixed η_1, κ_1 and η_2 we need to bound

$$\begin{aligned} & \sum_{\substack{I_1, I_2 \text{ good} \\ \ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) > 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}}} \sum_{J_1 \text{ good}} \left\langle \sum_{\substack{J_2 \\ 2^r \ell(J_1) < \ell(J_2) \leq 2^\ell}} \sum_{\kappa_2} \langle g_{\text{good}}, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle u_{J_2, \kappa_2} \right\rangle_{J_1} \\ & \quad \times \langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes 1 \rangle. \end{aligned}$$

We write $\langle g_{\text{good}}, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle = \langle g_{\text{good}}^{I_2, \eta_2}, u_{J_2, \kappa_2} \rangle$, where $g_{\text{good}}^{I_2, \eta_2}(y) = \langle g_{\text{good}}, h_{I_2, \eta_2} \rangle_1(y) = \int_{\mathbb{R}^n} g_{\text{good}}(x, y) h_{I_2, \eta_2}(x) d\mu_n(x)$. Then we note that

$$\sum_{\kappa_2} \langle g_{\text{good}}^{I_2, \eta_2}, u_{J_2, \kappa_2} \rangle u_{J_2, \kappa_2}(y) = \begin{cases} \Delta_{J_2}(g_{\text{good}}^{I_2, \eta_2})(y), & \text{if } \ell(J_2) < 2^\ell, \\ \Delta_{J_2}(g_{\text{good}}^{I_2, \eta_2})(y) + E_{J_2}(g_{\text{good}}^{I_2, \eta_2})(y), & \text{if } \ell(J_2) = 2^\ell. \end{cases}$$

Therefore, the sum inside the average over J_1 equals the constant $\langle g_{\text{good}}^{I_2, \eta_2} \rangle_{S(J_1)}$, where $S(J_1) \in \mathcal{D}'_m$ is the unique cube for which $\ell(S(J_1)) = 2^r \ell(J_1)$ and $J_1 \subset S(J_1)$. Note that for this collapse it is important that J_2 is not restricted to good cubes. The cube $S(J_1)$ exists since J_1 is good.

We are left to bound

$$\sum_{\eta_1, \kappa_1, \eta_2} \sum_{\substack{I_1, I_2 \text{ good} \\ \ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) > 2\ell(I_1)^\gamma \ell(I_2)^{1-\gamma_n}}} \sum_{J_2} \sum_{\substack{J_1 \text{ good} \\ J_1 \subset J_2 \\ \ell(J_1) = 2^{-r} \ell(J_2)}} \langle g_{\text{good}} \rangle_{J_2} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes 1 \rangle \\ \times \langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle.$$

We note that this equals

$$\left\langle \sum_{\eta_1, \kappa_1, \eta_2} \sum_{\substack{I_1, I_2 \text{ good} \\ \ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) > 2\ell(I_1)^\gamma \ell(I_2)^{1-\gamma_n}}} h_{I_2, \eta_2} \otimes (\Pi_{b_{I_1 I_2}^{\eta_1 \eta_2}}^{\kappa_1})^* f_{I_1}^{\eta_1}, g_{\text{good}} \right\rangle,$$

where $f_{I_1}^{\eta_1} = \langle f, h_{I_1, \eta_1} \rangle_1$, $b_{I_1 I_2}^{\eta_1 \eta_2} = \langle T^*(h_{I_2, \eta_2} \otimes 1), h_{I_1, \eta_1} \rangle_1$, and

$$\Pi_a^{\kappa_1} w = \sum_{J_2 \in \mathcal{D}'_m} \sum_{\substack{J_1 \in \mathcal{D}_{m, \text{good}} \\ J_1 \subset J_2 \\ \ell(J_1) = 2^{-r} \ell(J_2)}} \langle w \rangle_{J_2} \langle a, u_{J_1, \kappa_1} \rangle u_{J_1, \kappa_1}, \quad \kappa_1 \neq 0.$$

Here, of course, the pairings and averages are taken with respect to the μ_m measure.

There holds

$$\begin{aligned} & \left\| \left\langle \sum_{\eta_1, \kappa_1, \eta_2} \sum_{\substack{I_1, I_2 \text{ good} \\ \ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) > 2\ell(I_1)^\gamma \ell(I_2)^{1-\gamma_n}}} h_{I_2, \eta_2} \otimes (\Pi_{b_{I_1 I_2}^{\eta_1 \eta_2}}^{\kappa_1})^* f_{I_1}^{\eta_1}, g_{\text{good}} \right\rangle \right\|_{L^2(\mu)} \\ & \leq \sum_{\eta_1, \kappa_1, \eta_2} \left\| \sum_{\substack{I_1, I_2 \text{ good} \\ \ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) > 2\ell(I_1)^\gamma \ell(I_2)^{1-\gamma_n}}} h_{I_2, \eta_2} \otimes (\Pi_{b_{I_1 I_2}^{\eta_1 \eta_2}}^{\kappa_1})^* f_{I_1}^{\eta_1} \right\|_{L^2(\mu)} \|g\|_{L^2(\mu)} \\ & = \sum_{\eta_1, \kappa_1, \eta_2} \left(\sum_{I_2 \text{ good}} \left\| \sum_{\substack{I_1 \text{ good} \\ \ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) > 2\ell(I_1)^\gamma \ell(I_2)^{1-\gamma_n}}} (\Pi_{b_{I_1 I_2}^{\eta_1 \eta_2}}^{\kappa_1})^* f_{I_1}^{\eta_1} \right\|_{L^2(\mu_m)}^2 \right)^{1/2} \|g\|_{L^2(\mu)} \\ & \lesssim \sum_{\eta_1, \kappa_1, \eta_2} \left(\sum_{I_2} \left[\sum_{\substack{I_1: \ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) > 2\ell(I_1)^\gamma \ell(I_2)^{1-\gamma_n}}} \|b_{I_1 I_2}^{\eta_1 \eta_2}\|_{\text{BMO}_3^2(\mu_m)} \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)} \right]^2 \right)^{1/2} \|g\|_{L^2(\mu)}, \end{aligned}$$

where we used the orthonormality of the functions $h_{I_2, \eta_2} \in L^2(\mu_n)$, and the following paraproduct boundedness result:

4.6. Lemma. *For any $M > 1$, we have*

$$\|\Pi_a^{\kappa}\|_{L^2(\mu_m) \rightarrow L^2(\mu_m)} \lesssim \|a\|_{\text{BMO}_M^2(\mu_m)}.$$

The BMO norms above are estimated as follows:

4.7. Lemma. *If $\eta_1 \neq 0$, $\ell(I_1) \leq \ell(I_2)$ and $d(I_1, I_2) > 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}$, then there holds that*

$$\|b_{I_1 I_2}^{\eta_1 \eta_2}\|_{\text{BMO}_3^2(\mu_m)} \lesssim A_{I_1 I_2}^{\text{sep}}$$

for $b_{I_1 I_2}^{\eta_1 \eta_2} = \langle T^*(h_{I_2, \eta_2} \otimes 1), h_{I_1, \eta_1} \rangle_1$.

Proof. We fix a cube $V \subset \mathbb{R}^m$ and a function a such that $\text{spt } a \subset V$ and $\int a \, d\mu_m = 0$. We need to show that

$$\begin{aligned} & |\langle T(h_{I_1, \eta_1} \otimes a), h_{I_2, \eta_2} \otimes \chi_{3V} \rangle| \\ & + |\langle T(h_{I_1, \eta_1} \otimes a), h_{I_2, \eta_2} \otimes \chi_{(3V)^c} \rangle| \lesssim A_{I_1 I_2}^{\text{sep}} \|a\|_{L^2(\mu_m)} \mu_m(3V)^{1/2}. \end{aligned}$$

There holds with an arbitrary $z \in I_1$ that

$$\begin{aligned} & |\langle T(h_{I_1, \eta_1} \otimes a), h_{I_2, \eta_2} \otimes \chi_{3V} \rangle| \\ & = \left| \int_{I_1} \int_{I_2} [K_{a, \chi_{3V}}(x_1, y_1) - K_{a, \chi_{3V}}(x_1, z)] h_{I_1, \eta_1}(y_1) h_{I_2, \eta_2}(x_1) \, d\mu_n(x_1) \, d\mu_n(y_1) \right|, \end{aligned}$$

and this implies that

$$|\langle T(h_{I_1, \eta_1} \otimes a), h_{I_2, \eta_2} \otimes \chi_{3V} \rangle| \lesssim A_{I_1 I_2}^{\text{sep}} C(a, \chi_{3V}) \lesssim A_{I_1 I_2}^{\text{sep}} \|a\|_{L^2(\mu_m)} \mu_m(3V)^{1/2}.$$

An easy consequence of the Hölder estimates of the kernel K is that

$$\begin{aligned} & |\langle T(h_{I_1, \eta_1} \otimes a), h_{I_2, \eta_2} \otimes \chi_{(3V)^c} \rangle| \\ & \lesssim A_{I_1 I_2}^{\text{sep}} \int_V \int_{(3V)^c} \frac{\ell(V)^\beta}{|x_2 - c_V|^\beta \lambda_m(c_V, |x_2 - c_V|)} |a(y_2)| \, d\mu_m(x_2) \, d\mu_m(y_2) \\ & \lesssim A_{I_1 I_2}^{\text{sep}} \|a\|_{L^2(\mu_m)} \mu_m(V)^{1/2} \cdot \ell(V)^\beta \int_{(3V)^c} \frac{|x_2 - c_V|^{-\beta}}{\lambda_m(c_V, |x_2 - c_V|)} \, d\mu_m(x_2) \\ & \lesssim A_{I_1 I_2}^{\text{sep}} \|a\|_{L^2(\mu_m)} \mu_m(V)^{1/2}. \end{aligned}$$

This completes the proof of the lemma. □

Thus

$$\begin{aligned} & \left(\sum_{I_2} \left[\sum_{\substack{I_1: \ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) > 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}}} \|b_{I_1 I_2}^{\eta_1 \eta_2}\|_{\text{BMO}_3^2(\mu_m)} \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)} \right]^2 \right)^{1/2} \\ & \lesssim \left(\sum_{I_2} \left[\sum_{I_1: \ell(I_1) \leq \ell(I_2)} A_{I_1 I_2}^{\text{sep}} \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)} \right]^2 \right)^{1/2} \\ & \lesssim \left(\sum_{I_1} \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)}^2 \right)^{1/2} \leq \|f\|_{L^2(\mu)}, \end{aligned}$$

where we used Proposition 4.2 in the second-to-last step, and the orthonormality of the functions $h_{I_1}^{\eta_1}$, implicit in the notation $f_{I_1}^{\eta_1} = \langle f, h_{I_1}^{\eta_1} \rangle_1$, in the last step.

4.3. Separated/adjacent. Unlike in the full diagonal, no surgery is needed in this part of the summation. Simply by using the kernel $K_{u_{J_1, \kappa_1}, u_{J_2, \kappa_2}}$ and the knowledge that $C(u_{J_1, \kappa_1}, u_{J_2, \kappa_2}) \lesssim \|u_{J_1, \kappa_1}\|_{L^2(\mu_m)} \|u_{J_2, \kappa_2}\|_{L^2(\mu_m)} = 1$, we see like before that separation in \mathbb{R}^n gives

$$|\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle| \lesssim A_{I_1 I_2}^{\text{sep}}.$$

For a given J_1 there are only boundedly many J_2 for which $2^{-r}\ell(J_2) \leq \ell(J_1) \leq \ell(J_2)$ and $d(J_1, J_2) \leq 2\ell(J_1)^{\gamma_m} \ell(J_2)^{1-\gamma_m}$. Combining this fact with Hölder's inequality and Proposition 4.2 readily yields that the corresponding summation is dominated by $\|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)}$.

5. ADJACENT CUBES

We proceed to the case where the pair (I_1, I_2) consists of adjacent cubes. Since all separated cases were already handled, this leaves us with the two possibilities for the other pair (J_1, J_2) of being either nested or adjacent.

5.1. Surgery for adjacent cubes. For the analysis of the adjacent pairs (I_1, I_2) handled in this section, we will need to carry out a further splitting introduced by Nazarov, Treil and Volberg under the name *surgery*. We introduce it here in a pleasantly compact form.

Consider two adjacent cubes I_1, I_2 with $\ell(I_1) \leq \ell(I_2)$. Let $\theta \in (0, 1)$. We perform surgery on (I_1, I_2) with parameter $\theta > 0$. Let $j(\theta) \in \mathbb{Z}$ be such that $2^{-21}\theta \leq 2^{j(\theta)} < 2^{-20}\theta$. Let \mathcal{D}_n^* be yet another random grid in \mathbb{R}^n , independent of all other grids considered. Let $G := \{g \in \mathcal{D}_n^* : \ell(g) = 2^{j(\theta)}\ell(I_1)\}$, and for $x \in \mathbb{R}^n$, let $G(x)$ be the unique cube in G that contains x . We define

$$I_{1, \partial} := \{x \in I_1 : d(G(x), \partial I_2) < \theta\ell(I_2)/2\} \cup \{x \in I_1 \cap I_2 : d(x, \partial G(x)) < \theta\ell(G(x))\}.$$

Thus points in $I_{1, \partial}$ belong to I_1 , and are either close to the boundary of I_2 , or to the boundary of the grid G . The set $I_{1, \partial}$ depends on the set I_2 as well. However, we have

$$(5.1) \quad I_{1, \partial} \subset I_{1, \text{bad}} := I_1 \cap \left[\bigcup_{\substack{I'_2 \in \mathcal{D}'_n \\ \ell(I_1) \leq \ell(I'_2) \leq 2^r \ell(I_1)}} \{x : d(x, \partial I'_2) < \theta\ell(I'_2)\} \right. \\ \left. \cup \bigcup_{\substack{g \in \mathcal{D}_n^* \\ \ell(g) = 2^{j(\theta)} \ell(I_1)}} \{x : d(x, \partial g) < \theta\ell(g)\} \right],$$

which depends only on I_1 and the grids \mathcal{D}'_n and \mathcal{D}_n^* .

We set

$$I_{1, \text{sep}} := I_1 \setminus (I_{1, \partial} \cup I_2),$$

the part of I_1 strictly separated from I_2 . Finally, we have

$$I_{1, \Delta} := I_1 \setminus (I_{1, \partial} \cup I_{1, \text{sep}}) = \bigcup_i L_i,$$

where each L_i is of the form $L_i = (1 - \theta)g \cap I_1 \cap I_2$ for some $g \in G$, and $\#i \lesssim_\theta 1$. In fact, L_i is of the form $L_i = (1 - \theta)g$ unless it is close to the boundary of I_1 ; it cannot be close to the boundary of I_2 , since such cubes were already subtracted in the $I_{1,\partial}$ component.

We have the partition

$$I_1 = I_{1,\text{sep}} \cup I_{1,\partial} \cup I_{1,\Delta} = I_{1,\text{sep}} \cup I_{1,\partial} \cup \bigcup_i L_i,$$

and in a completely analogous manner also

$$I_2 = I_{2,\text{sep}} \cup I_{2,\partial} \cup I_{2,\Delta} = I_{2,\text{sep}} \cup I_{2,\partial} \cup \bigcup_j L_j.$$

A key observation is that all $L_i \subset I_1 \cap I_2$ appearing in the first union are cubes (of the form $(1 - \theta)g$ for $g \in G$) unless they are close to ∂I_1 , and they are never close to ∂I_2 , while the L_j in the second union are cubes unless they are close to ∂I_2 , and they are never close to ∂I_1 . Thus, all $L_i = L_j$ that appear in both unions are cubes and then $5L_i \subset I_1 \cap I_2$.

5.2. Adjacent/nested. The structure of the decomposition is the same as in the separated/nested case. Namely, we write $\langle T(h_{I_1,\eta_1} \otimes u_{J_1,\kappa_1}), h_{I_2,\eta_2} \otimes u_{J_2,\kappa_2} \rangle$ as the sum

$$\begin{aligned} & [\langle T(h_{I_1,\eta_1} \otimes u_{J_1,\kappa_1}), h_{I_2,\eta_2} \otimes u_{J_2,\kappa_2} \rangle - \langle u_{J_2,\kappa_2} \rangle_{J_1} \langle T(h_{I_1,\eta_1} \otimes u_{J_1,\kappa_1}), h_{I_2,\eta_2} \otimes 1 \rangle] \\ & + \langle u_{J_2,\kappa_2} \rangle_{J_1} \langle T(h_{I_1,\eta_1} \otimes u_{J_1,\kappa_1}), h_{I_2,\eta_2} \otimes 1 \rangle. \end{aligned}$$

Again, we start by dealing with the sum having the first term as the matrix element. Let $J_{2,1} \in \text{ch}(J_2)$ be such that $J_1 \subset J_{2,1}$. Note that $\langle u_{J_2,\kappa_2} \rangle_{J_1} = \langle u_{J_2,\kappa_2} \rangle_{J_{2,1}}$. Therefore, we have that

$$\langle T(h_{I_1,\eta_1} \otimes u_{J_1,\kappa_1}), h_{I_2,\eta_2} \otimes u_{J_2,\kappa_2} \rangle - \langle u_{J_2,\kappa_2} \rangle_{J_1} \langle T(h_{I_1,\eta_1} \otimes u_{J_1,\kappa_1}), h_{I_2,\eta_2} \otimes 1 \rangle$$

equals

$$\begin{aligned} & - \langle u_{J_2,\kappa_2} \rangle_{J_{2,1}} \langle T(h_{I_1,\eta_1} \otimes u_{J_1,\kappa_1}), h_{I_2,\eta_2} \otimes \chi_{J_{2,1}^c} \rangle \\ & + \sum_{\substack{J'_2 \in \text{ch}(J_2) \\ J'_2 \subset J_2 \setminus J_{2,1}}} \langle T(h_{I_1,\eta_1} \otimes u_{J_1,\kappa_1}), h_{I_2,\eta_2} \otimes u_{J_2,\kappa_2} \chi_{J'_2} \rangle. \end{aligned}$$

As J_1 is good, $J_1 \subset J_{2,1}$ and $\ell(J_1) \leq 2^{-r} \ell(J_{2,1})$, we have

$$d(J_1, J_{2,1}^c) \geq 4\ell(J_1)^{\gamma_m} \ell(J_{2,1})^{1-\gamma_m} > 2\ell(J_1)^{\gamma_m} \ell(J_2)^{1-\gamma_m}.$$

We state the following analogs of Lemma 4.3 and Lemma 4.4:

5.2. Lemma. *There holds*

$$|\langle u_{J_2,\kappa_2} \rangle_{J_{2,1}} \langle T(h_{I_1,\eta_1} \otimes u_{J_1,\kappa_1}), h_{I_2,\eta_2} \otimes \chi_{J_{2,1}^c} \rangle| \lesssim A_{J_1 J_2}^{\text{in}}.$$

5.3. Lemma. *If $J'_2 \in \text{ch}(J_2)$ and $J'_2 \subset J_2 \setminus J_{2,1}$, then*

$$|\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \chi_{J'_2} \rangle| \lesssim A_{J_1, J_2}^{\text{in}}.$$

Indeed, the proofs are similar as before except one uses the kernel $K_{h_{I_1, \eta_1}, h_{I_2, \eta_2}}$ and the fact that $C(h_{I_1, \eta_1}, h_{I_2, \eta_2}) \lesssim 1$. Recalling Lemma 4.5 and the fact that for a given I_1 there are only boundedly many I_2 for which $2^{-r} \ell(I_2) \leq \ell(I_1) \leq \ell(I_2)$ and $d(I_1, I_2) \leq 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}$, it is immediate that the corresponding series is dominated by $\|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)}$.

Therefore, we need only to consider the same summation but with the matrix element

$$\langle u_{J_2, \kappa_2} \rangle_{J_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes 1 \rangle.$$

Exactly as we have seen before, the series with this matrix element collapses to

$$\begin{aligned} & \left\langle \sum_{\eta_1, \kappa_1, \eta_2} \sum_{\substack{I_1, I_2 \text{ good} \\ 2^{-r} \ell(I_2) \leq \ell(I_1) \leq \ell(I_2) \\ d(I_1, I_2) \leq 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}}} h_{I_2, \eta_2} \otimes (\Pi_{b_{I_1 I_2}^{\kappa_1 \eta_1}})^* f_{I_1}^{\eta_1}, g_{\text{good}} \right\rangle \\ &= \sum_{\eta_1, \kappa_1, \eta_2} \sum_{\substack{I_1, I_2 \text{ good} \\ I_1 \sim I_2}} \langle (\Pi_{b_{I_1 I_2}^{\kappa_1 \eta_1}})^* f_{I_1}^{\eta_1}, g_{\text{good}, I_2}^{\eta_2} \rangle_2, \end{aligned}$$

where $f_{I_1}^{\eta_1} = \langle f, h_{I_1, \eta_1} \rangle_1$, $g_{\text{good}, I_2}^{\eta_2} = \langle g_{\text{good}}, h_{I_2, \eta_2} \rangle_1$, and $b_{I_1 I_2}^{\eta_1 \eta_2} = \langle T^*(h_{I_2, \eta_2} \otimes 1), h_{I_1, \eta_1} \rangle_1$. The summing condition tying I_1 and I_2 in the first line has been abbreviated to $I_1 \sim I_2$ in the second.

We can then estimate

$$\begin{aligned} & \left| \sum_{\substack{I_1, I_2 \text{ good} \\ I_1 \sim I_2}} \langle (\Pi_{b_{I_1 I_2}^{\kappa_1 \eta_1}})^* f_{I_1}^{\eta_1}, g_{\text{good}, I_2}^{\eta_2} \rangle_2 \right| \\ & \lesssim \sum_{\substack{I_1, I_2 \\ I_1 \sim I_2}} \|b_{I_1 I_2}^{\eta_1 \eta_2}\|_{\text{BMO}_{15}^2(\mu_m)} \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)} \|g_{I_2}^{\eta_2}\|_{L^2(\mu_m)}, \end{aligned}$$

where we also used the simple bound that $\|g_{\text{good}, I_2}^{\eta_2}\|_{L^2(\mu_m)} \leq \|g_{I_2}^{\eta_2}\|_{L^2(\mu_m)}$, which follows at once from the Haar expansions.

5.4. Lemma. *There holds that*

$$\|b_{I_1 I_2}^{\eta_1 \eta_2}\|_{\text{BMO}_{15}^2(\mu_m)} \leq C(\theta) + C\|T\|(H_{I_1, \text{bad}}^{\eta_1} + H_{I_2, \text{bad}}^{\eta_2})$$

where

$$H_{I_i, \text{bad}}^{\eta_i} = \left(\sum_{I'_i \in \text{ch}(I_i)} |\langle h_{I_i, \eta_i} \rangle_{I'_i}|^2 \mu_n(I'_{i, \text{bad}}) \right)^{1/2}$$

and $I_{i, \text{bad}}$ is defined as in (5.1) (that is, we use surgery with parameter $\theta > 0$).

Proof. We fix a cube $V \subset \mathbb{R}^m$ and a function a such that $\text{spt } a \subset V$ and $\int a \, d\mu_m = 0$. We should show that

$$|\langle T(h_{I_1, \eta_1} \otimes a), h_{I_2, \eta_2} \otimes 1 \rangle| \leq [C(\theta) + C\|T\|(H_{I_1, \text{bad}}^{\eta_1} + H_{I_2, \text{bad}}^{\eta_2})]\|a\|_{L^2(\mu_m)}\mu_m(15V)^{1/2}.$$

We split $1 = \chi_{3V} + \chi_{(3V)^c}$. The Hölder property of $K_{h_{I_1, \eta_1}, h_{I_2, \eta_2}}$ gives that

$$|\langle T(h_{I_1, \eta_1} \otimes a), h_{I_2, \eta_2} \otimes \chi_{(3V)^c} \rangle| \lesssim \|a\|_{L^2(\mu_m)}\mu_m(V)^{1/2}.$$

For the remaining part with χ_{3V} in place of 1, writing the Haar functions h_{I_i, η_i} as linear combinations of $\chi_{I'_i}$, where $I'_i \in \text{ch}(I_i)$, we are reduced to bounding

$$\sum_{I_1 \in \text{ch}(I_1)} \sum_{I_2 \in \text{ch}(I_2)} |\langle h_{I_1, \eta_1} \rangle_{I'_1}| |\langle h_{I_2, \eta_2} \rangle_{I'_2}| |\langle T(\chi_{I'_1} \otimes a), \chi_{I'_2} \otimes \chi_{3V} \rangle|.$$

We concentrate on one such pairing

$$\langle T(\chi_{I_1} \otimes a), \chi_{I_2} \otimes \chi_{3V} \rangle,$$

where we have dropped the primes from I'_i for brevity. This will be handled by applying the surgery to the pair (I_1, I_2) .

We write

$$\begin{aligned} \langle T(\chi_{I_1} \otimes a), \chi_{I_2} \otimes \chi_{3V} \rangle &= \sum_{\alpha \in \{\text{sep}, \partial\}} \langle T(\chi_{I_1, \alpha} \otimes a), \chi_{I_2} \otimes \chi_{3V} \rangle \\ &\quad + \sum_{\beta \in \{\text{sep}, \partial\}} \langle T(\chi_{I_1, \Delta} \otimes a), \chi_{I_2, \beta} \otimes \chi_{3V} \rangle \\ &\quad + \sum_{i \neq j} \langle T(\chi_{L_i} \otimes a), \chi_{L_j} \otimes \chi_{3V} \rangle \\ &\quad + \sum_{i=j} \langle T(\chi_{L_i} \otimes a), \chi_{L_i} \otimes \chi_{3V} \rangle. \end{aligned}$$

If $\alpha = \text{sep}$ or $\beta = \text{sep}$ or $i \neq j$, then the corresponding pairing is seen to be dominated by

$$C(\theta)\mu_n(I_1)^{1/2}\mu_n(I_2)^{1/2}\|a\|_{L^2(\mu_m)}\mu_m(3V)^{1/2}$$

using the size estimate of the kernel $K_{a, \chi_{3V}}$ together with the fact that the sets are separated by $c(\theta)\ell(I_1) \sim c(\theta)\ell(I_2)$. In the case $i \neq j$ a further large dependence on θ is gained from the summation $\sum_{i \neq j} 1$.

If $i = j$, then by the diagonal BMO assumptions the corresponding pairing is dominated by

$$C\mu_n(5L_i)\|a\|_{L^2(\mu_m)}\mu_m(15V)^{1/2} \leq C\mu_n(I_1)^{1/2}\mu_n(I_2)^{1/2}\|a\|_{L^2(\mu_m)}\mu_m(15V)^{1/2}.$$

A factor $C(\theta)$ is gained from the summation $\sum_{i=j} 1$.

Finally, the sum of the cases $\alpha = \partial$ and $\beta = \partial$ is dominated by

$$\|T\|(\mu_n(I_{1, \text{bad}})^{1/2}\mu_n(I_2)^{1/2} + \mu_n(I_1)^{1/2}\mu_n(I_{2, \text{bad}})^{1/2})\|a\|_{L^2(\mu_m)}\mu_m(3V)^{1/2}.$$

The BMO estimate now readily follows. \square

Recalling that

$$\sum_{I_1 \in \mathcal{D}_n} \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)}^2 \leq \|f\|_{L^2(\mu)}^2, \quad \sum_{I_2 \in \mathcal{D}'_n} \|g_{I_2}^{\eta_2}\|_{L^2(\mu_m)}^2 \leq \|g\|_{L^2(\mu)}^2,$$

we now know that

$$\begin{aligned} & \sum_{I_2 \in \mathcal{D}'_n} \sum_{\substack{I_1 \in \mathcal{D}_n \\ I_1 \sim I_2}} \|b_{I_1 I_2}^{\eta_1 \eta_2}\|_{\text{BMO}_{15}^2(\mu_m)} \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)} \|g_{I_2}^{\eta_2}\|_{L^2(\mu_m)} \\ & \leq \sum_{I_2 \in \mathcal{D}'_n} \sum_{\substack{I_1 \in \mathcal{D}_n \\ I_1 \sim I_2}} (C(\theta) + C\|T\|(H_{I_1, \text{bad}}^{\eta_1} + H_{I_2, \text{bad}}^{\eta_2})) \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)} \|g_{I_2}^{\eta_2}\|_{L^2(\mu_m)} \\ & \leq C(\theta) \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)} + C\|T\| \left[\sum_{I_1 \in \mathcal{D}_n} (H_{I_1, \text{bad}}^{\eta_1})^2 \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)}^2 \right]^{1/2} \|g\|_{L^2(\mu)} \\ & \quad + C\|T\| \|f\|_{L^2(\mu)} \left[\sum_{I_2 \in \mathcal{D}'_n} (H_{I_2, \text{bad}}^{\eta_2})^2 \|g_{I_2}^{\eta_2}\|_{L^2(\mu_m)}^2 \right]^{1/2}. \end{aligned}$$

The last two terms are symmetric to each other. For example, the first one can be estimated by

$$\begin{aligned} E_{w'_n} E_{w_n^*} [(H_{I_1, \text{bad}}^{\eta_1})^2] &= \sum_{I'_1 \in \text{ch}(I_1)} |\langle h_{I_1, \eta_1} \rangle_{I'_1}|^2 E_{w'_n} E_{w_n^*} [\mu_n(I'_{1, \text{bad}})] \\ &\leq c(\theta) \sum_{I'_1 \in \text{ch}(I_1)} |\langle h_{I_1, \eta_1} \rangle_{I'_1}|^2 \mu_n(I'_1) \leq c(\theta), \end{aligned}$$

where $c(\theta) \rightarrow 0$, when $\theta \rightarrow 0$. Furthermore, we have

$$\begin{aligned} & E_{w'_n} E_{w_n^*} \left[\sum_{I_1 \in \mathcal{D}_n} (H_{I_1, \text{bad}}^{\eta_1})^2 \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)}^2 \right]^{1/2} \\ & \leq \left[\sum_{I_1 \in \mathcal{D}_n} E_{w'_n} E_{w_n^*} [(H_{I_1, \text{bad}}^{\eta_1})^2] \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)}^2 \right]^{1/2} \\ & \leq c(\theta) \left[\sum_{I_1 \in \mathcal{D}_n} \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)}^2 \right]^{1/2} \leq c(\theta) \|f\|_{L^2(\mu)}. \end{aligned}$$

We have proved that the expectation of the series with the matrix element

$$\langle u_{J_2, \kappa_2} \rangle_{J_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes 1 \rangle$$

is dominated by $C(\theta) \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)} + c(\theta) \|T\| \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)}$. Fixing θ small we establish the bound $C\|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)} + 2^{-100} \|T\| \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)}$.

5.3. Adjacent/adjacent. We need to estimate the pairing

$$\begin{aligned} & \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle \\ &= \sum_{I'_1, I'_2, J'_1, J'_2} \langle h_{I_1, \eta_1} \rangle_{I'_1} \langle u_{J_1, \kappa_1} \rangle_{J'_1} \langle h_{I_2, \eta_2} \rangle_{I'_2} \langle u_{J_2, \kappa_2} \rangle_{J'_2} \langle T(\chi_{I'_1} \otimes \chi_{J'_1}), \chi_{I'_2} \otimes \chi_{J'_2} \rangle, \end{aligned}$$

where the summation is over all dyadic children I'_1 of I_1 , I'_2 of I_2 etc.

5.5. Lemma. For adjacent pairs (I_1, I_2) and (J_1, J_2) , we have

$$\begin{aligned} & |\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle| \\ & \leq C(\epsilon) + C\|T\| (H_{I_1, \text{bad}}^{\eta_1} + H_{I_2, \text{bad}}^{\eta_2} + H_{J_1, \text{bad}}^{\kappa_1} + H_{J_2, \text{bad}}^{\kappa_2}), \end{aligned}$$

where $\epsilon > 0$ is the surgery parameter both on \mathbb{R}^n and \mathbb{R}^m .

Proof. In the surgery of (J_1, J_2) we write K_i for the sets that correspond to the sets L_i in the surgery of (I_1, I_2) . We now decompose

$$\begin{aligned} \langle T(\chi_{I_1} \otimes \chi_{J_1}), \chi_{I_2} \otimes \chi_{J_2} \rangle &= \sum_{\alpha_1 \in \{\text{sep}, \partial\}} \langle T(\chi_{I_1, \alpha_1} \otimes \chi_{J_1}), \chi_{I_2} \otimes \chi_{J_2} \rangle \\ &+ \sum_{\beta_1 \in \{\text{sep}, \partial\}} \langle T(\chi_{I_1, \Delta} \otimes \chi_{J_1}), \chi_{I_2, \beta_1} \otimes \chi_{J_2} \rangle \\ &+ \sum_{\alpha_2 \in \{\text{sep}, \partial\}} \langle T(\chi_{I_1, \Delta} \otimes \chi_{J_1, \alpha_2}), \chi_{I_2, \Delta} \otimes \chi_{J_2} \rangle \\ &+ \sum_{\beta_2 \in \{\text{sep}, \partial\}} \langle T(\chi_{I_1, \Delta} \otimes \chi_{J_1, \Delta}), \chi_{I_2, \Delta} \otimes \chi_{J_2, \beta_2} \rangle \\ &+ \sum_{i_1 \neq j_1} \langle T(\chi_{L_{i_1}} \otimes \chi_{J_1, \Delta}), \chi_{L_{j_1}} \otimes \chi_{J_2, \Delta} \rangle \\ &+ \sum_{i_1=j_1} \sum_{i_2 \neq j_2} \langle T(\chi_{L_{i_1}} \otimes \chi_{K_{i_2}}), \chi_{L_{i_1}} \otimes \chi_{K_{j_2}} \rangle \\ &+ \sum_{i_1=j_1} \sum_{i_2=j_2} \langle T(\chi_{L_{i_1}} \otimes \chi_{K_{i_2}}), \chi_{L_{i_1}} \otimes \chi_{K_{i_2}} \rangle. \end{aligned}$$

If $i_1 = j_1$ and $i_2 = j_2$, then the weak boundedness property gives that the corresponding pairing is dominated by

$$C\mu_n(5L_{i_1})\mu_m(5K_{i_2}) \leq C\mu_n(I_1)^{1/2}\mu_n(I_2)^{1/2}\mu_m(J_1)^{1/2}\mu_m(J_2)^{1/2}.$$

A factor $C(\epsilon)$ is picked up from the summation $\sum_{i_1=j_1} \sum_{i_2=j_2} 1$.

The sum of the cases $\alpha_1 = \partial$ and $\beta_1 = \partial$ is dominated by

$$\|T\|(\mu_n(I_{1, \text{bad}})^{1/2}\mu_n(I_2)^{1/2} + \mu_n(I_1)^{1/2}\mu_n(I_{2, \text{bad}})^{1/2})\mu_m(J_1)^{1/2}\mu_m(J_2)^{1/2}.$$

The sum of the cases $\alpha_2 = \partial$ and $\beta_2 = \partial$ is dominated by

$$\|T\|\mu_n(I_1)^{1/2}\mu_n(I_2)^{1/2}(\mu_m(J_{1, \text{bad}})^{1/2}\mu_m(J_2)^{1/2} + \mu_m(J_1)^{1/2}\mu_m(J_{2, \text{bad}})^{1/2}).$$

In all the other cases we have separation in \mathbb{R}^n by $\epsilon\ell(I_1) \sim \epsilon\ell(I_2)$ or separation in \mathbb{R}^m by $\epsilon\ell(J_1) \sim \epsilon\ell(J_2)$. Partial kernel representations give the bound

$$C(\epsilon)\mu_n(I_1)^{1/2}\mu_n(I_2)^{1/2}\mu_m(J_1)^{1/2}\mu_m(J_2)^{1/2}.$$

We have shown that

$$\begin{aligned} & |\langle T(\chi_{I_1} \otimes \chi_{J_1}), \chi_{I_2} \otimes \chi_{J_2} \rangle| \\ & \leq C(\epsilon)\mu_n(I_1)^{1/2}\mu_n(I_2)^{1/2}\mu_m(J_1)^{1/2}\mu_m(J_2)^{1/2} \\ & + \|T\|(\mu_n(I_{1,\text{bad}})^{1/2}\mu_n(I_2)^{1/2} + \mu_n(I_1)^{1/2}\mu_n(I_{2,\text{bad}})^{1/2})\mu_m(J_1)^{1/2}\mu_m(J_2)^{1/2} \\ & + \|T\|\mu_n(I_1)^{1/2}\mu_n(I_2)^{1/2}(\mu_m(J_{1,\text{bad}})^{1/2}\mu_m(J_2)^{1/2} + \mu_m(J_1)^{1/2}\mu_m(J_{2,\text{bad}})^{1/2}). \end{aligned}$$

Apply this to $|\langle T(\chi_{I'_1} \otimes \chi_{J'_1}), \chi_{I'_2} \otimes \chi_{J'_2} \rangle|$, where $I'_1 \in \text{ch}(I_1)$, $I'_2 \in \text{ch}(I_2)$, $J'_1 \in \text{ch}(J_1)$ and $J'_2 \in \text{ch}(J_2)$ to get the claim. \square

With this lemma at hand, we see in the same way as in the adjacent/nested case that

$$\begin{aligned} E \Big| \sum_{I_1, I_2: I_1 \sim I_2} \sum_{J_1, J_2: J_1 \sim J_2} \langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle \langle g, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle \\ \times \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle \Big| \end{aligned}$$

is dominated by $C(\epsilon)\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)} + c(\epsilon)\|T\|\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}$, where $c(\epsilon) \rightarrow 0$ when $\epsilon \rightarrow 0$. Again, fixing ϵ small we establish the bound $C\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)} + 2^{-100}\|T\|\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}$.

6. NESTED CUBES

The only case left to deal with consists of quadruples of cubes such that $I_1 \subset I_{2,1} \in \text{ch}(I_2)$ and $J_1 \subset J_{2,1} \in \text{ch}(J_2)$. We then split

$$\begin{aligned} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle &= \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), s_{I_1 I_2}^{\eta_2} \otimes s_{J_1 J_2}^{\kappa_2} \rangle \\ &+ \langle h_{I_2, \eta_2} \rangle_{I_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), 1 \otimes s_{J_1 J_2}^{\kappa_2} \rangle \\ &+ \langle u_{J_2, \kappa_2} \rangle_{J_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), s_{I_1 I_2}^{\eta_2} \otimes 1 \rangle \\ &+ \langle h_{I_2, \eta_2} \rangle_{I_1} \langle u_{J_2, \kappa_2} \rangle_{J_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), 1 \rangle, \end{aligned}$$

where

$$s_{I_1 I_2}^{\eta_2} = \chi_{I_{2,1}^c}(h_{I_2, \eta_2} - \langle h_{I_2, \eta_2} \rangle_{I_{2,1}}) = -\langle h_{I_2, \eta_2} \rangle_{I_{2,1}} \chi_{I_{2,1}^c} + \sum_{\substack{I'_2 \in \text{ch}(I_2) \\ I'_2 \subset I_2 \setminus I_{2,1}}} h_{I_2, \eta_2} \chi_{I'_2}$$

and

$$s_{J_1 J_2}^{\kappa_2} = \chi_{J_{2,1}^c}(u_{J_2, \kappa_2} - \langle u_{J_2, \kappa_2} \rangle_{J_{2,1}}) = -\langle u_{J_2, \kappa_2} \rangle_{J_{2,1}} \chi_{J_{2,1}^c} + \sum_{\substack{J'_2 \in \text{ch}(J_2) \\ J'_2 \subset J_2 \setminus J_{2,1}}} u_{J_2, \kappa_2} \chi_{J'_2}.$$

6.1. Terms involving separation. We deal with the first three terms in the splitting, where at least one of $s_{I_1 I_2}^{\eta_2}$ or $s_{J_1 J_2}^{\kappa_2}$ appears.

We first consider the sum having the matrix element $\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), s_{I_1 I_2}^{\eta_2} \otimes s_{J_1 J_2}^{\kappa_2} \rangle$.

6.1. Lemma. *There holds that*

$$|\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), s_{I_1 I_2}^{\eta_2} \otimes s_{J_1 J_2}^{\kappa_2} \rangle| \lesssim A_{I_1 I_2}^{\text{in}} A_{J_1 J_2}^{\text{in}}.$$

Proof. We write out $\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), s_{I_1 I_2}^{\eta_2} \otimes s_{J_1 J_2}^{\kappa_2} \rangle$ as the following sum of four terms:

$$\begin{aligned} & \langle h_{I_2, \eta_2} \rangle_{I_{2,1}} \langle u_{J_2, \kappa_2} \rangle_{J_{2,1}} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), \chi_{I_{2,1}^c} \otimes \chi_{J_{2,1}^c} \rangle \\ & - \langle h_{I_2, \eta_2} \rangle_{I_{2,1}} \sum_{\substack{J'_2 \in \text{ch}(J_2) \\ J'_2 \subset J_2 \setminus J_{2,1}}} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), \chi_{I_{2,1}^c} \otimes u_{J_2, \kappa_2} \chi_{J'_2} \rangle \\ & - \langle u_{J_2, \kappa_2} \rangle_{J_{2,1}} \sum_{\substack{I'_2 \in \text{ch}(I_2) \\ I'_2 \subset I_2 \setminus I_{2,1}}} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \chi_{I'_2} \otimes \chi_{J_{2,1}^c} \rangle \\ & + \sum_{\substack{I'_2 \in \text{ch}(I_2) \\ I'_2 \subset I_2 \setminus I_{2,1}}} \sum_{\substack{J'_2 \in \text{ch}(J_2) \\ J'_2 \subset J_2 \setminus J_{2,1}}} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), h_{I_2, \eta_2} \chi_{I'_2} \otimes u_{J_2, \kappa_2} \chi_{J'_2} \rangle. \end{aligned}$$

We handle the second term here (the rest are bounded using the same tools). We record that

$$d(I_1, I_{2,1}^c) \geq 4\ell(I_1)^{\gamma_n} \ell(I_{2,1})^{1-\gamma_n} > 2\ell(I_1)^{\gamma_n} \ell(I_2)^{1-\gamma_n}$$

and

$$\begin{aligned} d(J_1, J'_2) & \geq d(J_1, J_{2,1}^c) \geq 4\ell(J_1)^{\gamma_m} \ell(J_{2,1})^{1-\gamma_m} \\ & = 4\ell(J_1)^{\gamma_m} \ell(J'_2)^{1-\gamma_m} > 2\ell(J_1)^{\gamma_m} \ell(J_2)^{1-\gamma_m}. \end{aligned}$$

Recall that $\eta_1 \neq 0$ and $\kappa_1 \neq 0$ (that is, the corresponding Haar functions have zero means), since for example I_1 is (much) smaller than I_2 and so $\ell(I_1) < 2^\ell$. Thus, with any $z \in I_1$ and $w \in J_1$ we may write

$$\begin{aligned} & \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), \chi_{I_{2,1}^c} \otimes u_{J_2, \kappa_2} \chi_{J'_2} \rangle \\ & = \int_{I_1} \int_{J_1} \int_{I_{2,1}^c} \int_{J'_2} [K(x, y) - K(x, (y_1, w)) - K(x, (z, y_2)) + K(x, (z, w))] \\ & \quad \times h_{I_1, \eta_1}(y_1) u_{J_1, \kappa_1}(y_2) u_{J_2, \kappa_2}(x_2) d\mu_m(x_2) d\mu_n(x_1) d\mu_m(y_2) d\mu_n(y_1). \end{aligned}$$

We get the bound

$$\begin{aligned} & |\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), \chi_{I_{2,1}^c} \otimes u_{J_2, \kappa_2} \chi_{J_2'} \rangle | \\ & \lesssim \mu_n(I_1)^{1/2} \ell(I_1)^\alpha \int_{\mathbb{R}^n \setminus B(z, d(I_1, I_{2,1}^c))} \frac{|x_1 - z|^{-\alpha}}{\lambda_n(z, |x_1 - z|)} d\mu_n(x_1) \\ & \quad \times \mu_m(J_1)^{1/2} \mu_m(J_2')^{1/2} \frac{\ell(J_1)^\beta}{d(J_1, J_2')^\beta \lambda_m(w, d(J_1, J_2'))}. \end{aligned}$$

As in the proof of Lemma 4.1 we get that

$$\begin{aligned} \frac{\ell(J_1)^\beta}{d(J_1, J_2')^\beta \lambda_m(w, d(J_1, J_2'))} & \lesssim \frac{\ell(J_1)^{\beta/2} \ell(J_2')^{\beta/2}}{D(J_1, J_2')^\beta \lambda_m(w, D(J_1, J_2'))} \\ & \leq \frac{\ell(J_1)^{\beta/2}}{\ell(J_2')^{\beta/2} \lambda_m(w, \ell(J_2'))}. \end{aligned}$$

As we have also previously seen, there holds that

$$\int_{\mathbb{R}^n \setminus B(z, d(I_1, I_{2,1}^c))} \frac{|x_1 - z|^{-\alpha}}{\lambda_n(z, |x_1 - z|)} d\mu_n(x_1) \lesssim d(I_1, I_{2,1}^c)^{-\alpha} \lesssim \ell(I_1)^{-\alpha/2} \ell(I_2)^{-\alpha/2}$$

and

$$\frac{\mu_m(J_2')^{1/2}}{\lambda_m(w, \ell(J_2'))} \lesssim \mu_m(J_{2,1})^{-1/2}.$$

It remains to recall that $|\langle h_{I_2, \eta_2} \rangle_{I_{2,1}}| \leq \mu_n(I_{2,1})^{-1/2}$. Indeed, putting these estimates together readily yields the bound $A_{I_1 I_2}^{\text{in}} A_{J_1 J_2}^{\text{in}}$. \square

Combining the previous lemma with Lemma 4.5 immediately gives that the summation with the matrix element $\langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), s_{I_1 I_2}^{\eta_2} \otimes s_{J_1 J_2}^{\kappa_2} \rangle$ is dominated by $\|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)}$.

We now deal with the sum having the matrix element $\langle u_{J_2, \kappa_2} \rangle_{J_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), s_{I_1 I_2}^{\eta_2} \otimes 1 \rangle$. Just like in the separated/nested case, we find out that the corresponding sum collapses to

$$\left\langle \sum_{\eta_1, \kappa_1, \eta_2} \sum_{\substack{I_1, I_2 \text{ good} \\ \ell(I_1) < 2^{-r} \ell(I_2) \\ I_1 \subset I_2}} h_{I_2, \eta_2} \otimes (\Pi_{b_{I_1 I_2}^{\kappa_1 \eta_2}})^* f_{I_1}^{\eta_1}, g_{\text{good}} \right\rangle,$$

where $f_{I_1}^{\eta_1} = \langle f, h_{I_1, \eta_1} \rangle_1$ and $b_{I_1 I_2}^{\eta_1 \eta_2} = \langle T^*(s_{I_1 I_2}^{\eta_2} \otimes 1), h_{I_1, \eta_1} \rangle_1$. Thus, it is enough to fix $\eta_1 \neq 0$, $\kappa_1 \neq 0$ and η_2 , and then show that

$$\left\| \sum_{\substack{I_1, I_2 \text{ good} \\ \ell(I_1) < 2^{-r} \ell(I_2) \\ I_1 \subset I_2}} h_{I_2, \eta_2} \otimes (\Pi_{b_{I_1 I_2}^{\kappa_1 \eta_2}})^* f_{I_1}^{\eta_1} \right\|_{L^2(\mu)} \lesssim \|f\|_{L^2(\mu)}.$$

An estimate of by now familiar nature shows that $\|b_{I_1 I_2}^{\eta_1 \eta_2}\|_{\text{BMO}_3^2(\mu_m)} \lesssim A_{I_1 I_2}^{\text{in}}$. We have by orthonormality and Lemma 4.5 that

$$\begin{aligned} & \left\| \sum_{\substack{I_1, I_2 \text{ good} \\ \ell(I_1) < 2^{-r} \ell(I_2) \\ I_1 \subset I_2}} h_{I_2, \eta_2} \otimes (\Pi_{b_{I_1 I_2}^{\eta_1 \eta_2}}^{\kappa_1})^* f_{I_1}^{\eta_1} \right\|_{L^2(\mu)} \\ & \lesssim \left(\sum_{I_2 \text{ good}} \left[\sum_{\substack{I_1 \text{ good} \\ \ell(I_1) < 2^{-r} \ell(I_2) \\ I_1 \subset I_2}} A_{I_1 I_2}^{\text{in}} \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)} \right]^2 \right)^{1/2} \\ & \lesssim \left(\sum_{I_1} \|f_{I_1}^{\eta_1}\|_{L^2(\mu_m)}^2 \right)^{1/2} \leq \|f\|_{L^2(\mu)}. \end{aligned}$$

Note that the proof of the boundedness of the summation with the matrix element $\langle h_{I_2, \eta_2} \rangle_{I_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), 1 \otimes s_{J_1 J_2}^{\kappa_2} \rangle$ is analogous.

6.2. Full paraproducts. We are reduced to considering the summation with the matrix element $\langle h_{I_2, \eta_2} \rangle_{I_1} \langle u_{J_2, \kappa_2} \rangle_{J_1} \langle T(h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}), 1 \rangle$. We fix $\eta_1 \neq 0$ and $\kappa_1 \neq 0$. We need to simplify the summation

$$\begin{aligned} & \sum_{\eta_2, \kappa_2} \sum_{I_1 \in \mathcal{D}_{n, \text{good}}} \sum_{J_1 \in \mathcal{D}_{m, \text{good}}} \sum_{\substack{I_2 \in \mathcal{D}'_n \\ 2^r \ell(I_1) < \ell(I_2) \leq 2^\ell}} \sum_{\substack{J_2 \in \mathcal{D}'_m \\ 2^r \ell(J_1) < \ell(J_2) \leq 2^\ell}} \langle h_{I_2, \eta_2} \rangle_{I_1} \langle u_{J_2, \kappa_2} \rangle_{J_1} \\ & \quad \times \langle g_{\text{good}}, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle \langle T^* 1, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle \langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle. \end{aligned}$$

Notice that

$$\left\langle \sum_{\substack{J_2 \in \mathcal{D}'_m \\ 2^r \ell(J_1) < \ell(J_2) \leq 2^\ell}} \sum_{\kappa_2} \langle g_{\text{good}}, h_{I_2, \eta_2} \otimes u_{J_2, \kappa_2} \rangle u_{J_2, \kappa_2} \right\rangle_{J_1} = \langle \langle g_{\text{good}}, h_{I_2, \eta_2} \rangle_1 \rangle_{S(J_1)},$$

where $S(J_1) \in \mathcal{D}'_m$ is the unique cube for which $\ell(S(J_1)) = 2^r \ell(J_1)$ and $J_1 \subset S(J_1)$. Continue to notice that

$$\left\langle \sum_{\substack{I_2 \in \mathcal{D}'_n \\ 2^r \ell(I_1) < \ell(I_2) \leq 2^\ell}} \sum_{\eta_2} h_{I_2, \eta_2} \otimes \langle g_{\text{good}}, h_{I_2, \eta_2} \rangle_1 \right\rangle_{I_1 \times S(J_1)} = \langle g_{\text{good}} \rangle_{S(I_1) \times S(J_1)},$$

where $S(I_1) \in \mathcal{D}'_n$ is the unique cube for which $\ell(S(I_1)) = 2^r \ell(I_1)$ and $I_1 \subset S(I_1)$. Therefore, the expression we started with collapses to

$$\begin{aligned} & \sum_{I_1 \in \mathcal{D}_{n,\text{good}}} \sum_{J_1 \in \mathcal{D}_{m,\text{good}}} \langle g_{\text{good}} \rangle_{S(I_1) \times S(J_1)} \langle T^* 1, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle \langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle \\ &= \sum_{I_2 \in \mathcal{D}'_n} \sum_{\substack{I_1 \in \mathcal{D}_{n,\text{good}} \\ I_1 \subset I_2 \\ \ell(I_1) = 2^{-r} \ell(I_2)}} \sum_{J_2 \in \mathcal{D}'_m} \sum_{\substack{J_1 \in \mathcal{D}_{m,\text{good}} \\ J_1 \subset J_2 \\ \ell(J_1) = 2^{-r} \ell(J_2)}} \langle g_{\text{good}} \rangle_{I_2 \times J_2} \langle T^* 1, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle \\ & \quad \times \langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle. \end{aligned}$$

One can write this as a pairing $\langle f, \Pi_{T^*1}^{\eta_1 \kappa_1} g_{\text{good}} \rangle$, where

$$\Pi_b^{\eta_1 \kappa_1} u = \sum_{I_2 \in \mathcal{D}'_n} \sum_{\substack{I_1 \in \mathcal{D}_{n,\text{good}} \\ I_1 \subset I_2 \\ \ell(I_1) = 2^{-r} \ell(I_2)}} \sum_{J_2 \in \mathcal{D}'_m} \sum_{\substack{J_1 \in \mathcal{D}_{m,\text{good}} \\ J_1 \subset J_2 \\ \ell(J_1) = 2^{-r} \ell(J_2)}} \langle u \rangle_{I_2 \times J_2} \langle b, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}.$$

6.2. Proposition. *There holds that*

$$\|\Pi_b^{\eta_1 \kappa_1} u\|_{L^2(\mu)} \leq 4 \|b\|_{\text{BMO}_{\text{prod}}(\mu)} \|u\|_{L^2(\mu)}, \quad \eta_1, \kappa_1 \neq 0.$$

Proof. There holds that

$$\begin{aligned} \|\Pi_b^{\eta_1 \kappa_1} u\|_{L^2(\mu)}^2 &\leq \sum_{S \in \mathcal{D}'} \sum_{\substack{R=I \times J \in \mathcal{D} \text{ good}, R \subset S \\ \text{gen}(R) = \text{gen}(S) + (r, r)}} |\langle u \rangle_S|^2 |\langle b, h_{I, \eta_1} \otimes u_{J, \kappa_1} \rangle|^2 \\ &= 2 \int_0^\infty \sum_{\substack{S \in \mathcal{D}' \\ |\langle u \rangle_S| > t}} \sum_{\substack{R=I \times J \in \mathcal{D} \text{ good}, R \subset S \\ \text{gen}(R) = \text{gen}(S) + (r, r)}} |\langle b, h_{I, \eta_1} \otimes u_{J, \kappa_1} \rangle|^2 t \, dt \\ &\leq 2 \int_0^\infty \sum_{\substack{S \in \mathcal{D}' \\ S \subset \{M_{\mathcal{D}'} u > t\}}} \sum_{\substack{R=I \times J \in \mathcal{D} \text{ good}, R \subset S \\ \text{gen}(R) = \text{gen}(S) + (r, r)}} |\langle b, h_{I, \eta_1} \otimes u_{J, \kappa_1} \rangle|^2 t \, dt \\ &\leq \|b\|_{\text{BMO}_{\text{prod}}(\mu)}^2 2 \int_0^\infty \mu(\{M_{\mathcal{D}'} u > t\}) t \, dt \\ &= \|b\|_{\text{BMO}_{\text{prod}}(\mu)}^2 \|M_{\mathcal{D}'} u\|_{L^2(\mu)}^2 \\ &\leq 16 \|b\|_{\text{BMO}_{\text{prod}}(\mu)}^2 \|u\|_{L^2(\mu)}^2, \end{aligned}$$

which is what we wanted to prove. \square

7. MIXED PARAPRODUCTS

In the last few sections, we have completed the estimation of the part of the Haar series expansion of $\langle Tf, g \rangle$ involving quadruples of cubes with $\ell(I_1) \leq \ell(I_2)$ and $\ell(J_1) \leq \ell(J_2)$. The other three subseries with one or both of these inequalities reversed are completely symmetric, except for one detail: in the case that the two

inequalities go in different directions, we end up with a different form of the full paraproduct, which we call mixed.

The following explanation reveals how it appears. In the case $\ell(I_1) \leq \ell(I_2)$ and $\ell(J_1) > \ell(J_2)$ one needs to deal with a full paraproduct, which is different in an essential way:

$$\begin{aligned} \Pi_{\text{mixed}, b}^{\eta_1 \kappa_2} u := & \sum_{I_2 \in \mathcal{D}'_n} \sum_{\substack{I_1 \in \mathcal{D}_{n, \text{good}} \\ I_1 \subset I_2 \\ \ell(I_1) = 2^{-r} \ell(I_2)}} \sum_{J_1 \in \mathcal{D}_m} \sum_{\substack{J_2 \in \mathcal{D}'_{m, \text{good}} \\ J_2 \subset J_1 \\ \ell(J_2) = 2^{-r} \ell(J_1)}} \left\langle u, h_{I_1, \eta_1} \otimes \frac{\chi_{J_1}}{\mu_m(J_1)} \right\rangle \\ & \times \left\langle b, h_{I_1, \eta_1} \otimes u_{J_2, \kappa_2} \right\rangle \frac{\chi_{I_2}}{\mu_n(I_2)} \otimes u_{J_2, \kappa_2}, \end{aligned}$$

$\eta_1 \neq 0, \kappa_2 \neq 0$. In the actual summation one has $b = T_1(1)$. We need to separately demonstrate that $\|\Pi_{\text{mixed}, b}^{\eta_1 \kappa_2} u\|_{L^2(\mu)} \lesssim \|b\|_{\text{BMO}_{\text{prod}}(\mu)} \|u\|_{L^2(\mu)}$. Following [20], we dualize with a $v \in L^2(\mu)$ and write

$$\langle \Pi_{\text{mixed}, b}^{\eta_1 \eta_2} u, v \rangle = \langle b, \Lambda(u, v) \rangle$$

for a suitable bilinear form Λ . We then need the two estimates

$$|\langle b, f \rangle| \lesssim \|b\|_{\text{BMO}_{\text{prod}}(\mu)} \|f\|_{H^1(\mu)}, \quad \|\Lambda(u, v)\|_{H^1(\mu)} \lesssim \|u\|_{L^2(\mu)} \|v\|_{L^2(\mu)}$$

for a suitable (ad hoc) “ H^1 ” norm $\|\cdot\|_{H^1(\mu)}$.

7.1. H^1 -BMO type duality inequality. We denote $\mathcal{C}_K := \{I \in \mathcal{D}_n \text{ good} : I \subset K, \text{gen}(I) = \text{gen}(K) + r\}$ for $K \in \mathcal{D}'_n$, similarly $\mathcal{C}_L \subset \mathcal{D}_m$ for $L \in \mathcal{D}'_m$ and $\mathcal{C}_S := \mathcal{C}_K \times \mathcal{C}_L \subset \mathcal{D}$ for $S = K \times L \in \mathcal{D}' = \mathcal{D}'_n \times \mathcal{D}'_m$. Also, we denote $S(I) = K$ if $I \in \mathcal{C}_K$.

7.1. Definition. We define the square function

$$\begin{aligned} S_{\mathcal{D}'_n \mathcal{D}'_m}^{\mathcal{D}'_n \mathcal{D}'_m} f &= \left(\sum_{\substack{\eta_1 \neq 0 \\ \kappa_1 \neq 0}} \sum_{\substack{I_1 \in \mathcal{D}_{n, \text{good}} \\ J_1 \in \mathcal{D}_{m, \text{good}}}} |\langle f, h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} \rangle|^2 \frac{\chi_{S(I_1)} \otimes \chi_{S(J_1)}}{\mu_n(S(I_1)) \mu_m(S(J_1))} \right)^{1/2} \\ &= \left(\sum_{S \in \mathcal{D}'} \sum_{R \in \mathcal{C}_S} |\langle f, h_R \rangle|^2 \frac{\chi_S}{\mu(S)} \right)^{1/2}, \end{aligned}$$

where we suppressed the η_1, κ_1 summation and denoted $h_R = h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1}$.

The following proposition is the main result of this section.

7.2. Proposition. *If $b \in \text{BMO}_{\text{prod}}(\mu)$, then for all dyadic grids $\mathcal{D}_n, \mathcal{D}'_n, \mathcal{D}_m$ and \mathcal{D}'_m , and for all functions f of the form*

$$f = \sum_{\substack{\eta_1 \neq 0 \\ \kappa_1 \neq 0}} \sum_{\substack{I_1 \in \mathcal{D}_{n, \text{good}} \\ J_1 \in \mathcal{D}_{m, \text{good}}}} \lambda_{I_1 J_1}^{\eta_1 \kappa_1} h_{I_1, \eta_1} \otimes u_{J_1, \kappa_1} = \sum_{S \in \mathcal{D}'} \sum_{R \in \mathcal{C}_S} \lambda_R h_R$$

there holds that

$$|\langle b, f \rangle| \lesssim \|b\|_{\text{BMO}_{\text{prod}}(\mu)} \|S_{\mathcal{D}'_n \mathcal{D}'_m}^{\mathcal{D}'_n \mathcal{D}'_m} f\|_{L^1(\mu)}.$$

Here we use the same suppressed notation as in the above definition.

Proof. We begin by simply estimating

$$|\langle b, f \rangle| \leq \sum_{S \in \mathcal{D}'} \sum_{R \in \mathcal{C}_S} |\langle b, h_R \rangle \lambda_R|.$$

Let

$$(7.3) \quad \phi := S_{\mathcal{D}_n' \mathcal{D}_m'} f = \left(\sum_{S \in \mathcal{D}'} \sum_{R \in \mathcal{C}_S} |\lambda_R|^2 \frac{\chi_S}{\mu(S)} \right)^{1/2},$$

and set $\Omega_k := \{\phi > 2^k\}$, and $\mathcal{S}_k := \{S \in \mathcal{D}' : \mu(S \cap \Omega_k) > \frac{1}{2}\mu(S)\}$. Note that if $S \notin \bigcup_{k \in \mathbb{Z}} \mathcal{S}_k$, then $\mu(S \cap \{\phi = 0\}) \geq \frac{1}{2}\mu(S)$, and so

$$\sum_{R \in \mathcal{C}_S} |\lambda_R|^2 \leq 2 \sum_{R \in \mathcal{C}_S} |\lambda_R|^2 \frac{\mu(S \cap \{\phi = 0\})}{\mu(S)} \leq 2 \int_{\{\phi=0\}} \phi^2 d\mu = 0.$$

Therefore, all the relevant S in the sum of interest belong to at least one \mathcal{S}_k . Also, $\bigcap_{k \in \mathbb{Z}} \Omega_k = \{\phi = \infty\}$, which has measure zero if the right side of the claim is finite (which we may assume). This means that if $S \in \bigcap_{k \in \mathbb{Z}} \mathcal{S}_k$, then $\mu(S) = 0$ and so $\lambda_R = 0$ for every $R \in \mathcal{C}_S$.

We infer that for all the relevant S in the sum of interest there exist a unique $k_S \in \mathbb{Z}$ so that $S \in \mathcal{S}_{k_S} \setminus \mathcal{S}_{k_S+1}$. Thus, we may reorganize

$$\sum_{S \in \mathcal{D}'} \sum_{R \in \mathcal{C}_S} |\langle b, h_R \rangle \lambda_R| = \sum_{k \in \mathbb{Z}} \sum_{S \in \mathcal{S}_k \setminus \mathcal{S}_{k+1}} \sum_{R \in \mathcal{C}_S} |\langle b, h_R \rangle \lambda_R|.$$

We then bound

$$(7.4) \quad \begin{aligned} \sum_{S \in \mathcal{S}_k \setminus \mathcal{S}_{k+1}} \sum_{R \in \mathcal{C}_S} |\langle b, h_R \rangle \lambda_R| &\leq \left(\sum_{S \in \mathcal{S}_k \setminus \mathcal{S}_{k+1}} \sum_{R \in \mathcal{C}_S} |\langle b, h_R \rangle|^2 \right)^{1/2} \\ &\quad \times \left(\sum_{S \in \mathcal{S}_k \setminus \mathcal{S}_{k+1}} \sum_{R \in \mathcal{C}_S} |\lambda_R|^2 \right)^{1/2}. \end{aligned}$$

Clearly $S \in \mathcal{S}_k$ means that $S \subset \tilde{\Omega}_k := \{M\chi_{\Omega_k} > \frac{1}{2}\}$, where M is the strong dyadic maximal operator with respect to \mathcal{D}' . Thus

$$\begin{aligned} \left(\sum_{S \in \mathcal{S}_k \setminus \mathcal{S}_{k+1}} \sum_{R \in \mathcal{C}_S} |\langle b, h_R \rangle|^2 \right)^{1/2} &\leq \left(\sum_{S \subset \tilde{\Omega}_k} \sum_{R \in \mathcal{C}_S} |\langle b, h_R \rangle|^2 \right)^{1/2} \\ &\leq \|b\|_{\text{BMO}_{\text{prod}}(\mu)} \mu(\tilde{\Omega}_k)^{1/2} \lesssim \|b\|_{\text{BMO}_{\text{prod}}(\mu)} \mu(\Omega_k)^{1/2}. \end{aligned}$$

For the other factor in (7.4), we use the condition that $S \notin \mathcal{S}_{k+1}$ to write

$$1 \leq 2\mu(S \setminus \Omega_{k+1})/\mu(S),$$

and again the condition that $S \in \mathcal{S}_k$ to see that $S \subset \tilde{\Omega}_k$, so that

$$\begin{aligned} \sum_{S \in \mathcal{S}_k \setminus \mathcal{S}_{k+1}} \sum_{R \in \mathcal{C}_S} |\lambda_R|^2 &\leq 2 \int_{\tilde{\Omega}_k \setminus \Omega_{k+1}} \sum_{S \in \mathcal{S}_k \setminus \mathcal{S}_{k+1}} \sum_{R \in \mathcal{C}_S} |\lambda_R|^2 \frac{\chi_S}{\mu(S)} d\mu \\ &\leq 2 \int_{\tilde{\Omega}_k \setminus \Omega_{k+1}} \phi^2 d\mu \leq 2 \cdot (2^{k+1})^2 \mu(\tilde{\Omega}_k) \lesssim 2^{2k} \mu(\Omega_k). \end{aligned}$$

Substituting back to (7.4), we obtain

$$\sum_{S \in \mathcal{S}_k \setminus \mathcal{S}_{k+1}} \sum_{R \in \mathcal{C}_S} |\langle b, h_R \rangle \lambda_R| \lesssim \|b\|_{\text{BMO}_{\text{prod}}(\mu)} \mu(\Omega_k)^{1/2} \cdot 2^k \mu(\Omega_k)^{1/2},$$

and hence

$$\sum_{S \in \mathcal{D}'} \sum_{R \in \mathcal{C}_S} |\langle b, h_R \rangle \lambda_R| \lesssim \sum_{k \in \mathbb{Z}} \|b\|_{\text{BMO}_{\text{prod}}(\mu)} 2^k \mu(\Omega_k) \lesssim \|b\|_{\text{BMO}_{\text{prod}}(\mu)} \|\phi\|_{L^1(\mu)}. \quad \square$$

7.2. Boundedness of the mixed paraproduct. We are now ready to prove:

7.5. Proposition. *There holds that $\|\Pi_{\text{mixed}, b}^{\eta_1 \kappa_2} u\|_{L^2(\mu)} \lesssim \|b\|_{\text{BMO}_{\text{prod}}(\mu)} \|u\|_{L^2(\mu)}$.*

Proof. This proof is adapted from [20].

Let us abbreviate $L := \|b\|_{\text{BMO}_{\text{prod}}(\mu)}$. We will show that

$$|\langle \Pi_{\text{mixed}, b}^{\eta_1 \kappa_2} u, v \rangle| \lesssim L \|u\|_{L^2(\mu)} \|v\|_{L^2(\mu)}.$$

We write

$$\begin{aligned} \langle \Pi_{\text{mixed}, b}^{\eta_1 \kappa_2} u, v \rangle &= \left\langle b, \sum_{I_2 \in \mathcal{D}'_n} \sum_{\substack{I_1 \in \mathcal{D}_{n, \text{good}} \\ I_1 \subset I_2 \\ \ell(I_1) = 2^{-r} \ell(I_2)}} \sum_{J_1 \in \mathcal{D}_m} \sum_{\substack{J_2 \in \mathcal{D}'_{m, \text{good}} \\ J_2 \subset J_1 \\ \ell(J_2) = 2^{-r} \ell(J_1)}} \left\langle u, h_{I_1, \eta_1} \otimes \frac{\chi_{J_1}}{\mu_m(J_1)} \right\rangle \right. \\ &\quad \left. \times \left\langle v, \frac{\chi_{I_2}}{\mu_n(I_2)} \otimes u_{J_2, \kappa_2} \right\rangle h_{I_1, \eta_1} \otimes u_{J_2, \kappa_2} \right\rangle. \end{aligned}$$

By Proposition 7.2, we have

$$|\langle \Pi_{\text{mixed}, b}^{\eta_1 \kappa_2} u, v \rangle| \lesssim L \|S_{\mathcal{D}_n \mathcal{D}'_m}^{\mathcal{D}'_n \mathcal{D}_m} f\|_{L^1(\mu)},$$

where

$$f = \sum_{I_1 \in \mathcal{D}_{n, \text{good}}} \sum_{J_2 \in \mathcal{D}'_{m, \text{good}}} \left\langle u, h_{I_1, \eta_1} \otimes \frac{\chi_{S(J_2)}}{\mu_m(S(J_2))} \right\rangle \left\langle v, \frac{\chi_{S(I_1)}}{\mu_n(S(I_1))} \otimes u_{J_2, \kappa_2} \right\rangle h_{I_1, \eta_1} \otimes u_{J_2, \kappa_2}.$$

We estimate $S_{\mathcal{D}_n \mathcal{D}'_m}^{\mathcal{D}'_n \mathcal{D}_m} f$ pointwise. Let us first note that

$$\begin{aligned} (S_{\mathcal{D}_n \mathcal{D}'_m}^{\mathcal{D}'_n \mathcal{D}_m} f)^2 &\leq \sum_{I_1 \in \mathcal{D}_{n, \text{good}}} \sum_{J_2 \in \mathcal{D}'_{m, \text{good}}} \left| \left\langle u, h_{I_1, \eta_1} \otimes \frac{\chi_{S(J_2)}}{\mu_m(S(J_2))} \right\rangle \right|^2 \\ &\quad \times \left| \left\langle v, \frac{\chi_{S(I_1)}}{\mu_n(S(I_1))} \otimes u_{J_2, \kappa_2} \right\rangle \right|^2 \frac{\chi_{S(I_1)} \otimes \chi_{S(J_2)}}{\mu_n(S(I_1)) \mu_m(S(J_2))}. \end{aligned}$$

This implies that

$$\begin{aligned} S_{\mathcal{D}_n \mathcal{D}'_m} f &\leq \left(\sum_{I_1 \in \mathcal{D}_{n,\text{good}}} \frac{\chi_{S(I_1)}}{\mu_n(S(I_1))} \otimes \sup_{J_2 \in \mathcal{D}'_{m,\text{good}}} \left[\left| \langle u, h_{I_1, \eta_1} \otimes \frac{\chi_{S(J_2)}}{\mu_m(S(J_2))} \rangle \right|^2 \chi_{S(J_2)} \right] \right)^{1/2} \\ &\quad \times \left(\sum_{J_2 \in \mathcal{D}'_{m,\text{good}}} \sup_{I_1 \in \mathcal{D}_{n,\text{good}}} \left[\left| \langle v, \frac{\chi_{S(I_1)}}{\mu_n(S(I_1))} \otimes u_{J_2, \kappa_2} \rangle \right|^2 \chi_{S(I_1)} \right] \otimes \frac{\chi_{S(J_2)}}{\mu_m(S(J_2))} \right)^{1/2}. \end{aligned}$$

This yields the final pointwise bound

$$\begin{aligned} S_{\mathcal{D}_n \mathcal{D}'_m} f &\leq \left(\sum_{I_1 \in \mathcal{D}_{n,\text{good}}} \frac{\chi_{S(I_1)}}{\mu_n(S(I_1))} \otimes M_{\mathcal{D}_m}(\langle u, h_{I_1, \eta_1} \rangle_1^2) \right)^{1/2} \\ &\quad \times \left(\sum_{J_2 \in \mathcal{D}'_{m,\text{good}}} M_{\mathcal{D}'_n}(\langle v, u_{J_2, \kappa_2} \rangle_2^2) \otimes \frac{\chi_{S(J_2)}}{\mu_m(S(J_2))} \right)^{1/2}. \end{aligned}$$

Using the above pointwise bound together with Cauchy–Schwarz we get that

$$\begin{aligned} \|S_{\mathcal{D}_n \mathcal{D}'_m} f\|_{L^1(\mu)} &\leq \left\| \sum_{I_1 \in \mathcal{D}_{n,\text{good}}} \frac{\chi_{S(I_1)}}{\mu_n(S(I_1))} \otimes M_{\mathcal{D}_m}(\langle u, h_{I_1, \eta_1} \rangle_1^2) \right\|_{L^1(\mu)}^{1/2} \\ &\quad \times \left\| \sum_{J_2 \in \mathcal{D}'_{m,\text{good}}} M_{\mathcal{D}'_n}(\langle v, u_{J_2, \kappa_2} \rangle_2^2) \otimes \frac{\chi_{S(J_2)}}{\mu_m(S(J_2))} \right\|_{L^1(\mu)}^{1/2}. \end{aligned}$$

The proof is ended by noting that

$$\begin{aligned} &\left\| \sum_{I_1 \in \mathcal{D}_{n,\text{good}}} \frac{\chi_{S(I_1)}}{\mu_n(S(I_1))} \otimes M_{\mathcal{D}_m}(\langle u, h_{I_1, \eta_1} \rangle_1^2) \right\|_{L^1(\mu)} \\ &= \sum_{I_1 \in \mathcal{D}_{n,\text{good}}} \int_{\mathbb{R}^m} M_{\mathcal{D}_m}(\langle u, h_{I_1, \eta_1} \rangle_1^2) d\mu_m \\ &\lesssim \sum_{I_1 \in \mathcal{D}_n} \int_{\mathbb{R}^m} |\langle u, h_{I_1, \eta_1} \rangle_1|^2 d\mu_m \lesssim \|u\|_{L^2(\mu)}^2. \end{aligned}$$

□

8. NECESSITY OF $T1 \in \text{BMO}_{\text{prod}}(\mu)$

8.1. Dyadic Journé’s lemma for general product measures. We prove Journé’s covering lemma [13] with general measures. In this section, we assume that $\mu = \mu_n \times \mu_m$ is an arbitrary product of a Borel measure μ_n in \mathbb{R}^n and μ_m in \mathbb{R}^m . That is, we don’t assume any growth conditions. Again, we have the dyadic systems \mathcal{D}_n and \mathcal{D}_m in \mathbb{R}^n and \mathbb{R}^m respectively.

For $R = I \times J \in \mathcal{D} := \mathcal{D}_n \times \mathcal{D}_m$, write $R^{(i,j)} := I^{(i)} \times J^{(j)}$ and

$$\text{gen}_1(R) := \text{gen}(I), \quad \text{gen}(R) := (\text{gen}(I), \text{gen}(J)),$$

where $I^{(i)}$ is the usual dyadic ancestor, and $\text{gen}(I)$ is the usual generation of a dyadic cube.

Let $\Omega \subset \mathbb{R}^{n+m}$ be such a set that $\mu(\Omega) < \infty$ and that for every $x \in \Omega$ there exists $R \in \mathcal{D}$ so that $x \in R \subset \Omega$. We let

$$\tilde{\Omega} := \{M\chi_{\Omega} > \tfrac{1}{2}\},$$

where M is the strong dyadic maximal operator with respect to \mathcal{D} . Define the embeddedness of R in Ω as

$$\text{emb}_1(R; \Omega) := \sup\{k : R^{(k,0)} \subset \tilde{\Omega}\}.$$

This notation and some related inspiration is derived from [3].

We say that a dyadic rectangle $R = I \times J \subset \Omega$ is 2-maximal if $I \times \tilde{J} \not\subset \Omega$ for any dyadic $\tilde{J} \supsetneq J$. (Obviously we could define $\text{emb}_2(R; \Omega)$ and 1-maximality analogously. The main point is that in the following result we need to consider embeddedness and maximality with respect to different variables.)

8.1. Theorem (Journé's lemma). *Let $\omega : \mathbb{N} \rightarrow \mathbb{R}_+$ be a decreasing function with the property that $\sum_{k=0}^{\infty} \omega(k) < \infty$. Then there holds that*

$$\sum_{\substack{R \subset \Omega \\ \text{2-maximal}}} \omega(\text{emb}_1(R; \Omega)) \times \mu(R) \leq 2 \sum_{k=0}^{\infty} \omega(k) \times \mu(\Omega).$$

Proof. Let us define $\delta(k) := \omega(k) - \omega(k+1) \geq 0$, so that $\omega(j) = \sum_{k=j}^{\infty} \delta(k)$. We have

$$\begin{aligned} \sum_{\substack{R \subset \Omega \\ \text{2-maximal}}} \omega(\text{emb}_1(R; \Omega)) \mu(R) &= \sum_{j=0}^{\infty} \omega(j) \sum_{\substack{R \subset \Omega \text{ 2-maximal} \\ \text{emb}_1(R; \Omega) = j}} \mu(R) \\ (8.2) \quad &= \sum_{k=0}^{\infty} \delta(k) \sum_{\substack{R \subset \Omega \text{ 2-maximal} \\ \text{emb}_1(R; \Omega) \leq k}} \mu(R) = \sum_{k=0}^{\infty} \delta(k) \sum_{i=0}^k \sum_{R \in \mathcal{R}(k,i)} \mu(R), \end{aligned}$$

where

$$\mathcal{R}(k, i) := \left\{ R \subset \Omega \text{ 2-maximal; } \text{emb}_1(R; \Omega) \leq k, \text{ gen}_1(R) \equiv i \pmod{k+1} \right\}.$$

We make the following observation about two intersecting $R = I \times J$ and $R' = I' \times J'$ in $\mathcal{R}(k, i)$. By definition, we have $I \cap I' \neq \emptyset$ and $J \cap J' \neq \emptyset$. Suppose first that $I = I'$. But then also $J = J'$, since otherwise one of R and R' could not be 2-maximal. Thus $I \neq I'$ whenever $R \neq R'$. Let us then consider the case that $I \subsetneq I'$. Then we must have $J \supseteq J'$, since otherwise $R = I \times J \subsetneq I \times J' \subset R' \subset \Omega$ would contradict the 2-maximality of R .

For every $R = I \times J \in \mathcal{R}(k, i)$, consider the subset

$$E(R) := R \setminus \bigcup_{\substack{R' = I' \times J' \in \mathcal{R}(k,i) \\ I' \supsetneq I}} R' = I \times \left(J \setminus \bigcup_{\substack{I' \times J' \in \mathcal{R}(k,i) \\ I' \supsetneq I}} J' \right),$$

where we observe that $J' \subset J$ for all relevant J' in the union (i.e., those for which R' intersects R), by what we just checked above. The sets $E(R)$, $R \in \mathcal{R}(k, i)$, are pairwise disjoint; namely, if two different $R = I \times J$, $R' = I' \times J' \in \mathcal{R}(k, i)$ intersect, the previous paragraph implies that either $I' \supsetneq I$, in which case $R' \cap E(R) = \emptyset$, or else $I \supsetneq I'$, in which case $R \cap E(R') = \emptyset$.

We claim that

$$(8.3) \quad \mu(E(R)) = \mu_n(I)\mu_m\left(J \setminus \bigcup_{\substack{I' \times J' \in \mathcal{R}(k, i) \\ I' \supsetneq I}} J'\right) \geq \frac{1}{2}\mu(R) = \frac{1}{2}\mu_n(I)\mu_m(J).$$

This is trivial if $\mu_n(I) = 0$. Otherwise, suppose that the opposite estimate is valid. Observe that $I \subsetneq I'$ and $\text{gen}(I) \equiv i \equiv \text{gen}(I') \pmod{k+1}$ imply that $I^{(k+1)} \subset I'$. Thus

$$R^{(k+1,0)} \setminus \Omega \subset R^{(k+1,0)} \setminus \bigcup_{\substack{R' = I' \times J' \in \mathcal{R}(k, i) \\ I' \supsetneq I}} R' = I^{(k+1)} \times \left(J \setminus \bigcup_{\substack{I' \times J' \in \mathcal{R}(k, i) \\ I' \supsetneq I}} J'\right).$$

Hence

$$\begin{aligned} \mu(R^{(k+1,0)} \setminus \Omega) &\leq \mu_n(I^{(k+1)})\mu_m\left(J \setminus \bigcup_{\substack{I' \times J' \in \mathcal{R}(k, i) \\ I' \supsetneq I}} J'\right) \\ &< \mu_n(I^{(k+1)}) \cdot \frac{1}{2}\mu_m(J) = \frac{1}{2}\mu(R^{(k+1,0)}) \end{aligned}$$

and therefore

$$\inf_{R^{(k+1,0)}} M\chi_\Omega \geq \frac{\mu(R^{(k+1,0)} \cap \Omega)}{\mu(R^{(k+1,0)})} > 1 - \frac{1}{2} = \frac{1}{2}.$$

But this means that $R^{(k+1,0)} \subset \tilde{\Omega}$, contradicting $\text{emb}_1(R; \Omega) \leq k$.

Hence the claim (8.3) is valid. But this means that

$$\sum_{R \in \mathcal{R}(k, i)} \mu(R) \leq \sum_{R \in \mathcal{R}(k, i)} 2\mu(E(R)) \leq 2\mu(\Omega),$$

where the last step is due to the fact that the sets $E(R)$ are pairwise disjoint and all contained in Ω . Substituting back to (8.2) shows that

$$\sum_{\substack{R \subset \Omega \\ 2\text{-maximal}}} \omega(\text{emb}_1(R; \Omega))\mu(R) \leq \sum_{k=0}^{\infty} \delta(k) \times (k+1) \times 2\mu(\Omega) = 2 \sum_{k=0}^{\infty} \omega(k) \times \mu(\Omega). \quad \square$$

8.2. A class of bi-parameter singular integrals. In this section we return to the setting of upper doubling measures. Let $\mu = \mu_n \times \mu_m$, where μ_n and μ_m are upper doubling measures on \mathbb{R}^n and \mathbb{R}^m respectively. The corresponding dominating functions are denoted by λ_n and λ_m . We consider what seems like a slightly restricted class of product Calderón–Zygmund operators than above. In fact, the definition will be close to the classical setting of Journé, just with general

measures. We want to demonstrate the necessity of the product BMO condition for this subclass.

We consider product Calderón–Zygmund operators $T \in \mathcal{L}(L^2(\mu))$ with the following vector-valued Calderón–Zygmund structure. When viewing T as an operator on $L^2(\mu_n; L^2(\mu_m))$, it has a kernel $T_1(x_1, y_1)$ with values in $L^2(\mu_m)$ -bounded Calderón–Zygmund operators, such that

$$\langle \Phi, T\Psi \rangle = \iint \langle \Phi(x_1), T_1(x_1, y_1)\Psi(y_1) \rangle_2 d\mu_n(x_1) d\mu_m(y_1)$$

whenever the $L^2(\mu_m)$ -valued functions Φ and Ψ are disjointly supported. This kernel is required to satisfy the standard estimates

$$\|T_1(x_1, y_1)\|_{CZ(\mu_m)} \lesssim \frac{1}{\lambda_n(x_1, |x_1 - y_1|)}$$

and

$$\|T_1(x_1, y_1) - T_1(x'_1, y_1)\|_{CZ(\mu_m)} \lesssim \left(\frac{|x_1 - x'_1|}{|x_1 - y_1|} \right)^\alpha \frac{1}{\lambda_n(x_1, |x_1 - y_1|)}$$

for $|x_1 - y_1| > 2|x_1 - x'_1|$. (We do not require regularity in the second variable for the present considerations.) Here $\|\cdot\|_{CZ(\mu_m)}$ designates the sum of the $\mathcal{L}(L^2(\mu_m))$ -norm of the operator, and the Calderón–Zygmund constants of its (scalar-valued) kernel. We impose the analogous conditions when viewing T as an operator on the space $L^2(\mu_m; L^2(\mu_n))$.

8.4. Remark. Let $\Phi = \Phi_1 \otimes \Phi_2$ and $\Psi = \Psi_1 \otimes \Psi_2$ with $\text{spt } \Phi_1 \cap \text{spt } \Psi_1 = \emptyset$. Then with our original notation we have $K_{\Phi_2, \Psi_2}(x_1, y_1) = \langle \Phi_2, T_1(x_1, y_1)\Psi_2 \rangle_2$, and we have the Calderón–Zygmund bounds with the constant $C(\Phi_2, \Psi_2) \lesssim \|\Phi_2\|_{L^2(\mu_m)}\|\Psi_2\|_{L^2(\mu_m)}$. If also $\text{spt } \Phi_2 \cap \text{spt } \Psi_2 = \emptyset$, then one has a representation with the full kernel $K(x, y) = K_{T_1(x_1, y_1)}(x_2, y_2)$. We will show that for this subclass of operators our new non-homogeneous product BMO condition is also necessary. Although, let us note again that for example the regularity in the second variable is not needed in this section. We use this subclass, since it seems that it allows more control in situations where not all of the functions are of tensor product type, and this type of control is needed here for the first time.

8.5. Theorem. *Let T be an $L^2(\mu)$ bounded product Calderón–Zygmund operator in the sense above, and $b \in L^\infty(\mu)$. Then*

$$\|Tb\|_{\text{BMO}_{\text{prod}}(\mu)} \lesssim \|b\|_{L^\infty(\mu)}.$$

8.3. Initial considerations. We may assume by homogeneity that $\|b\|_{L^\infty(\mu)} = 1$. By Definition 2.11 and Remark 2.12, we need to check the estimate

$$\sum_{\substack{S \in \mathcal{D}' \\ S \subset \Omega}} \sum_{\substack{R \in \mathcal{D} \text{ good, } R \subset S \\ \text{gen}(R) = \text{gen}(S) + (r, r)}} |\langle h_R, Tb \rangle|^2 \lesssim \mu(\Omega)$$

for any dyadic systems $\mathcal{D} = \mathcal{D}_n \times \mathcal{D}_m$ and $\mathcal{D}' = \mathcal{D}'_n \times \mathcal{D}'_m$, where $\Omega \subset \mathbb{R}^{n+m}$ is a bounded set such that $\mu(\Omega) < \infty$ and that for every $x \in \Omega$ there exists $S \in \mathcal{D}'$

so that $x \in S \subset \Omega$. We have also used the following short-hand notation: If $R = I \times J$, then $h_R = h_{I \times J} = h_I \otimes h_J$ is shorthand for $h_{I,\eta} \otimes u_{J,\kappa}$ for $\eta \neq 0$ and $\kappa \neq 0$ (that is, only cancellative Haar functions appear).

We start by setting up some further simplifying notation. For $K \in \mathcal{D}'_n$, let

$$\mathcal{C}_K := \{I \in \mathcal{D}_n \text{ good} : I \subset K, \text{gen}(I) = \text{gen}(K) + r\}$$

and

$$\hat{K} := \bigcup_{I \in \mathcal{C}_K} I \subset \{x_1 \in K : d(x_1, K^c) \geq 2^{-r\gamma_n} \ell(K)\},$$

where the last containment is immediate from the definition of goodness. We use similar notation \mathcal{C}_L and \hat{L} for $L \in \mathcal{D}'_m$.

We decompose $b = b\chi_{\tilde{\Omega}} + b\chi_{\tilde{\Omega}^c}$, where $\tilde{\Omega} := \{M_{\mathcal{D}'}\chi_{\Omega} > \frac{1}{2}\}$ and $M_{\mathcal{D}'}$ is the strong dyadic maximal operator with respect to \mathcal{D}' . Then

$$\begin{aligned} \sum_{S=K \times L \subset \Omega} \sum_{R \in \mathcal{C}_K \times \mathcal{C}_L} |\langle h_R, T(b\chi_{\tilde{\Omega}}) \rangle|^2 &\leq \sum_{R \in \mathcal{D}} |\langle h_R, T(b\chi_{\tilde{\Omega}}) \rangle|^2 \leq \|T(b\chi_{\tilde{\Omega}})\|_2^2 \\ &\lesssim \|b\chi_{\tilde{\Omega}}\|_2^2 \lesssim \mu(\tilde{\Omega}) \lesssim \mu(\Omega), \end{aligned}$$

using

$$\mu(\tilde{\Omega}) = \mu(M_{\mathcal{D}'}\chi_{\Omega} > \frac{1}{2}) \leq 4\|M_{\mathcal{D}'}\chi_{\Omega}\|_2^2 \lesssim \|\chi_{\Omega}\|_2^2 = \mu(\Omega)$$

in the last step.

So it suffices to prove the estimate with $b\chi_{\tilde{\Omega}^c}$ in place of b . For simplicity, we denote it again by b , assuming from now on that $\text{spt } b \subset \tilde{\Omega}^c$.

For every $L \in \mathcal{D}'_m$, let \mathcal{F}_L be the collection of the maximal $F \in \mathcal{D}'_n$ such that $F \times L \subset \tilde{\Omega}$, and denote $F_L := \bigcup_{F \in \mathcal{F}_L} F$. If there are any cubes $F \in \mathcal{D}'_n$ with $F \times L \subset \tilde{\Omega}$, then the maximal cubes exist by standard properties of dyadic cubes, except if there is a strictly increasing sequence of cubes $F_k \in \mathcal{D}'_n$ with $F_k \times L \subset \tilde{\Omega}$. Since our dyadic systems do not have quadrants, such a sequence will exhaust all of \mathbb{R}^n , and then in fact $\mathbb{R}^n \times L \subset \tilde{\Omega}$. In this case we set $\mathcal{F}_L := \{\mathbb{R}^n\}$, $F_L := \mathbb{R}^n$. We refer to this as the *degenerate case*.

These notions lead to the splitting $b = b_L^1 + b_L^2$, where

$$b_L^1(x_1, x_2) := b(x_1, x_2)\chi_{F_L}(x_1).$$

Note that if $b_L^1(y_1, y_2) \neq 0$, then $y_1 \in F_L$ while $(y_1, y_2) \in \tilde{\Omega}^c \subset (F_L \times L)^c$, and hence $y_2 \in L^c$. Thus $\text{spt } b_L^1 \subset F_L \times L^c$. Note that $b_L^1 = b$, $b_L^2 = 0$ in the degenerate case.

Now we estimate the left hand side in the Theorem by

$$\sum_{i=1}^2 \sum_{\substack{S=K \times L \in \mathcal{D}' \\ S \subset \Omega}} \sum_{\substack{I \in \mathcal{C}_K \\ J \in \mathcal{C}_L}} |\langle h_{I \times J}, T b_L^i \rangle|^2.$$

8.4. Analysis of b_L^1 . We do not need to pay any special attention to the possible degenerate cases in this analysis. The set F_L will appear every now and then, but it can equally well be the full space \mathbb{R}^n or a subset thereof. The only property that we need is that $\mu_n(F_L) < \infty$ whenever this set appears. To see this, observe that b_L^1 only appears in the pairing $\langle h_{I \times J}, T b_L^1 \rangle$, where $J \in \mathcal{C}_L$. If $\mu_m(L) = 0$, then $h_{I \times J} = 0$ for all $J \in \mathcal{C}_L$, and we can ignore such pairings. On the other hand, if $\mu_m(L) > 0$, then

$$\mu_n(F_L)\mu_m(L) = \mu(F_L \times L) \leq \mu(\tilde{\Omega}) < \infty$$

implies that $\mu_n(F_L) < \infty$, as claimed.

We begin with:

8.6. Lemma. *For a fixed $L \in \mathcal{D}'_m$, the sum over the other variables is estimated as*

$$\begin{aligned} & \sum_{K \in \mathcal{D}'_n} \sum_{\substack{I \in \mathcal{C}_K \\ J \in \mathcal{C}_L}} |\langle h_{I \times J}, T b_L^1 \rangle|^2 \\ & \lesssim \mu_m(\hat{L}) \int_{\mathbb{R}^n} \int_{\text{spt } b_L^1(y_1, \cdot)} \left(\frac{\ell(L)}{d(\hat{L}, y_2)} \right)^\beta \frac{d\mu_m(y_2)}{\lambda_m(x_{\hat{L}}, d(\hat{L}, y_2))} d\mu_n(y_1), \end{aligned}$$

where $x_{\hat{L}}$ is an arbitrarily fixed point in \hat{L} .

Note that we have dropped the restriction that $K \times L \subset \Omega$; this part of the estimate is true even if we allow K to range over all dyadic intervals.

Proof. We can write

$$\begin{aligned} \langle h_{I \times J}, T b_L^1 \rangle &= \iint h_J(x_2) \langle h_I, T_2(x_2, y_2) b_L^1(\cdot, y_2) \rangle_1 d\mu_m(x_2) d\mu_m(y_2) \\ &= \iint h_J(x_2) \langle h_I, [T_2(x_2, y_2) - T_2(c_J, y_2)] b_L^1(\cdot, y_2) \rangle_1 d\mu_m(x_2) d\mu_m(y_2), \end{aligned}$$

where $T_2(x_2, y_2)$ is the $\mathcal{L}(L^2(\mu_n))$ -valued kernel of T when interpreted as an operator on $L^2(\mu_m; L^2(\mu_n))$; it acts on the function $b_L^1(\cdot, y_2) \in L^2(\mu_n)$ (the function is bounded and supported on F_L), and $\langle \cdot, \cdot \rangle_1$ designates the inner product in $L^2(\mu_n)$. Hence, abbreviating

$$B_J(x_2, y_2) := [T_2(x_2, y_2) - T_2(c_J, y_2)] b_L^1(\cdot, y_2) \in L^2(\mu_n),$$

we have that

$$\langle h_{I \times J}, T b_L^1 \rangle = \iint h_J(x_2) \langle h_I, B_J(x_2, y_2) \rangle_1 d\mu_m(x_2) d\mu_m(y_2).$$

Thus, by the orthonormality of the Haar functions $h_I \in L^2(\mu_n)$, we have (with a fixed $J \in \mathcal{C}_L$) that

$$\begin{aligned} \sum_{\substack{K \in \mathcal{D}'_n \\ I \in \mathcal{C}_K}} |\langle h_{I \times J}, T b_L^1 \rangle|^2 &\leq \sum_{I \in \mathcal{D}_n} |\langle h_{I \times J}, T b_L^1 \rangle|^2 \\ &\leq \left[\iint |h_J(x_2)| \left(\sum_{I \in \mathcal{D}_n} |\langle h_I, B_J(x_2, y_2) \rangle_1|^2 \right)^{1/2} d\mu_m(x_2) d\mu_m(y_2) \right]^2 \\ &\leq \left[\iint |h_J(x_2)| \|B_J(x_2, y_2)\|_{L^2(\mu_n)} d\mu_m(x_2) d\mu_m(y_2) \right]^2. \end{aligned}$$

Since $x_2 \in J$, we have $|x_2 - c_J| \leq \ell(J)/2$. Moreover, since $y_2 \in L^c$, there holds that

$$|c_J - y_2| \geq d(J, L^c) \geq \ell(J)^{\gamma_m} \ell(L)^{1-\gamma_m} \geq \ell(J) \geq 2|x_2 - c_J|.$$

Hence the Hölder estimate for T_2 gives that

$$\|T_2(x_2, y_2) - T_2(c_J, y_2)\|_{\mathcal{L}(L^2(\mu_n))} \lesssim \left(\frac{\ell(J)}{|c_J - y_2|} \right)^\beta \frac{1}{\lambda_m(c_J, |c_J - y_2|)}.$$

It follows that

$$\begin{aligned} \sum_{\substack{K \in \mathcal{D}'_n \\ I \in \mathcal{C}_K \\ J \in \mathcal{C}_L}} |\langle h_{I \times J}, T b_L^1 \rangle|^2 &\lesssim \sum_{J \in \mathcal{C}_L} \left[\int |h_J(x_2)| \int \left(\frac{\ell(J)}{|c_J - y_2|} \right)^\beta \frac{\|b_L^1(\cdot, y_2)\|_{L^2(\mu_n)}}{\lambda_m(c_J, |c_J - y_2|)} d\mu_m(y_2) d\mu_m(x_2) \right]^2 \\ &\lesssim \sum_{J \in \mathcal{C}_L} \mu_m(J) \left[\int \left(\frac{\ell(J)}{|c_J - y_2|} \right)^\beta \frac{\|b_L^1(\cdot, y_2)\|_{L^2(\mu_n)}}{\lambda_m(c_J, |c_J - y_2|)} d\mu_m(y_2) \right]^2. \end{aligned}$$

By Cauchy–Schwarz, the last brackets squared are dominated by

$$\begin{aligned} &\left[\int_{J^c} \left(\frac{\ell(J)}{|c_J - y_2|} \right)^\beta \frac{1}{\lambda_m(c_J, |c_J - y_2|)} d\mu_m(y_2) \right] \\ &\quad \times \left[\int \left(\frac{\ell(J)}{|c_J - y_2|} \right)^\beta \frac{\|b_L^1(\cdot, y_2)\|_{L^2(\mu_n)}^2}{\lambda_m(c_J, |c_J - y_2|)} d\mu_m(y_2) \right]. \end{aligned}$$

It is a standard fact that

$$\int_{J^c} \left(\frac{\ell(J)}{|c_J - y_2|} \right)^\beta \frac{1}{\lambda_m(c_J, |c_J - y_2|)} d\mu_m(y_2) \lesssim 1.$$

To estimate the other bracket we need a few observations. Since $c_J \in J \in \mathcal{C}_L$, we have that $c_J \in \hat{L}$ and so $|c_J - y_2| \geq d(\hat{L}, y_2)$. So we may estimate

$$\left(\frac{\ell(J)}{|c_J - y_2|} \right)^\beta \frac{1}{\lambda_m(c_J, |c_J - y_2|)} \leq \left(\frac{\ell(L)}{d(\hat{L}, y_2)} \right)^\beta \frac{1}{\lambda_m(c_J, d(\hat{L}, y_2))}.$$

We still want to get rid of the last dependence on J , the c_J on $\lambda_m(c_J, d(\hat{L}, y_2))$. To this end, let $x_{\hat{L}}$ be an arbitrarily fixed point in \hat{L} as in the statement of the Lemma. Since $c_J, x_{\hat{L}} \in L$, we have $|c_J - x_{\hat{L}}| \leq \ell(L)$. Moreover, we have $d(\hat{L}, y_2) \geq d(\hat{L}, L^c) \geq 2^{-r\gamma_m} \ell(L) \geq 2^{-r\gamma_m} |c_J - x_{\hat{L}}|$. Therefore, we can estimate that

$$\lambda_m(c_J, d(\hat{L}, y_2)) \sim \lambda_m(c_J, 2^{r\gamma_m} d(\hat{L}, y_2)) \sim \lambda_m(x_{\hat{L}}, 2^{r\gamma_m} d(\hat{L}, y_2)) \sim \lambda_m(x_{\hat{L}}, d(\hat{L}, y_2)).$$

So for the other bracket we get the bound

$$\begin{aligned} & \int \left(\frac{\ell(J)}{|c_J - y_2|} \right)^\beta \frac{\|b_L^1(\cdot, y_2)\|_{L^2(\mu_n)}^2}{\lambda_m(c_J, |c_J - y_2|)} d\mu_m(y_2) \\ & \lesssim \int \left(\frac{\ell(L)}{d(\hat{L}, y_2)} \right)^\beta \frac{\|b_L^1(\cdot, y_2)\|_{L^2(\mu_n)}^2}{\lambda_m(x_{\hat{L}}, d(\hat{L}, y_2))} d\mu_m(y_2) \\ & \lesssim \int_{\mathbb{R}^n} \int_{\text{spt } b_L^1(y_1, \cdot)} \left(\frac{\ell(L)}{d(\hat{L}, y_2)} \right)^\beta \frac{d\mu_m(y_2)}{\lambda_m(x_{\hat{L}}, d(\hat{L}, y_2))} d\mu_n(y_1). \end{aligned}$$

Since these bounds do not depend on J , the proof is completed by observing that

$$\sum_{J \in \mathcal{C}_L} \mu_m(J) = \mu_m(\hat{L}). \quad \square$$

It remains to sum the bound of the Lemma over all $L \in \mathcal{D}'_m$. (Note that $F_L = \emptyset = \text{spt } b_L^1$ if there is no dyadic interval F such that $F \times L \subset \tilde{\Omega}$.) Thus we need to estimate

$$\begin{aligned} & \sum_{\substack{K \times L \in \mathcal{D}' \\ K \times L \subset \Omega}} \sum_{\substack{I \in \mathcal{C}_K \\ J \in \mathcal{C}_L}} |\langle h_{I \times J}, T b_L^1 \rangle|^2 \\ (8.7) \quad & \lesssim \sum_{L \in \mathcal{D}'_m} \mu_m(\hat{L}) \int_{\mathbb{R}^n} \int_{\text{spt } b_L^1(y_1, \cdot)} \left(\frac{\ell(L)}{d(\hat{L}, y_2)} \right)^\beta \frac{d\mu_m(y_2)}{\lambda_m(x_{\hat{L}}, d(\hat{L}, y_2))} d\mu_n(y_1) \\ & = \iint \sum_{L: \hat{L} \ni x_2} \int_{\text{spt } b_L^1(y_1, \cdot)} \left(\frac{\ell(L)}{d(\hat{L}, y_2)} \right)^\beta \frac{d\mu_m(y_2)}{\lambda_m(x_{\hat{L}}, d(\hat{L}, y_2))} d\mu(y_1, x_2), \end{aligned}$$

where the last equality follows by writing $\mu_m(\hat{L}) = \int \chi_{\hat{L}}(x_2) d\mu_m(x_2)$ and reorganizing.

Note that the outermost double integral may be restricted to $\tilde{\Omega}$. Indeed, let $x_2 \in L$ and $\text{spt } b_L^1(y_1, \cdot)$ be nontrivial. Then necessarily $y_1 \in F_L$, and thus $(y_1, x_2) \in F_L \times L \subset \tilde{\Omega}$.

For each fixed $(y_1, x_2) \in \tilde{\Omega}$, let $T = T(y_1, x_2)$ denote the maximal $L \in \mathcal{D}'_m$ appearing in (8.7) with the property that $x_2 \in \hat{L}$ and $\text{spt } b_L^1(y_1, \cdot) \neq \emptyset$, so in particular $y_1 \in F_L$. The existence of these maximal cubes L can be seen as follows: Since Ω is bounded, we can obviously add the implicit restriction $\ell(L) \leq \text{diam}(\Omega)$ to the summation on the first line, and then all other lines, in (8.7). Thus the size of

the relevant cubes L is bounded from above, and then the existence of maximal cubes is a standard property of dyadic grids.

By definition, we have $F_T \times T \subset \tilde{\Omega}$ and $y_1 \in F_T$. So in particular $\{y_1\} \times T \subset \tilde{\Omega}$. Since $\text{spt } b_L^1 \subset \tilde{\Omega}^c$, it follows that $\text{spt } b_L^1(y_1, \cdot) \subset T^c$, so any y_2 in the integral satisfies $y_2 \in \text{spt } b_L^1(y_1, \cdot) \subset T^c$.

If $z \in \hat{L}$, then there is a good cube $J \in \mathcal{D}_m$ so that $z \in J \subset L \subset T$ and $\ell(J) = 2^{-r}\ell(L) \leq 2^{-r}\ell(T)$. The goodness implies that

$$|z - y_2| \geq d(J, T^c) \geq \ell(J)^{\gamma_m} \ell(T)^{1-\gamma_m} = 2^{-r\gamma_m} \ell(L)^{\gamma_m} \ell(T)^{1-\gamma_m}.$$

We may conclude that every y_2 in the integral satisfies the bound $d(\hat{L}, y_2) \geq 2^{-r\gamma_m} \ell(L)^{\gamma_m} \ell(T)^{1-\gamma_m}$. So we may estimate

$$\begin{aligned} & \int_{\text{spt } b_L^1(y_1, \cdot)} \frac{d(\hat{L}, y_2)^{-\beta} d\mu_m(y_2)}{\lambda_m(x_{\hat{L}}, d(\hat{L}, y_2))} \\ & \leq \sum_{j=0}^{\infty} \int_{\{y_2: 2^j 2^{-r\gamma_m} \ell(L)^{\gamma_m} \ell(T)^{1-\gamma_m} \leq d(\hat{L}, y_2) < 2^{j+1} 2^{-r\gamma_m} \ell(L)^{\gamma_m} \ell(T)^{1-\gamma_m}\}} \frac{d(\hat{L}, y_2)^{-\beta} d\mu_m(y_2)}{\lambda_m(x_{\hat{L}}, d(\hat{L}, y_2))} \\ & \lesssim \sum_{j=0}^{\infty} (2^j 2^{-r\gamma_m} \ell(L)^{\gamma_m} \ell(T)^{1-\gamma_m})^{-\beta} \frac{\mu_m(B(x_{\hat{L}}, [2 + 2^{r\gamma_m}] \cdot 2^j 2^{-r\gamma_m} \ell(L)^{\gamma_m} \ell(T)^{1-\gamma_m}))}{\lambda_m(x_{\hat{L}}, 2^j 2^{-r\gamma_m} \ell(L)^{\gamma_m} \ell(T)^{1-\gamma_m})} \\ & \lesssim \ell(L)^{-\beta\gamma_m} \ell(T)^{-\beta(1-\gamma_m)}. \end{aligned}$$

This gives us the bound

$$\begin{aligned} \sum_{L: \hat{L} \ni x_2} \int_{\text{spt } b_L^1(y_1, \cdot)} \left(\frac{\ell(L)}{d(\hat{L}, y_2)} \right)^{\beta} \frac{d\mu_m(y_2)}{\lambda_m(x_{\hat{L}}, d(\hat{L}, y_2))} & \lesssim \sum_{L: x_2 \in L \subset T} \left(\frac{\ell(L)}{\ell(T)} \right)^{\beta(1-\gamma_m)} \\ & = \sum_{j=0}^{\infty} 2^{-\beta(1-\gamma_m)j} \lesssim 1. \end{aligned}$$

Substituting back to (8.7) and recalling that the outer double integral is restricted to $\tilde{\Omega}$, we deduce that

$$\sum_{K \times L \subset \Omega} \sum_{\substack{I \in \mathcal{C}_K \\ J \in \mathcal{C}_L}} |\langle h_{I \times J}, T b_L^1 \rangle|^2 \lesssim \iint_{\tilde{\Omega}} d\mu(y_1, x_2) = \mu(\tilde{\Omega}) \lesssim \mu(\Omega),$$

as required.

8.5. Analysis of b_L^2 . Here we sum the multiple series in a different order:

8.8. Lemma. *For a fixed $K \in \mathcal{D}'_n$, the sum over the other variables is estimated as*

$$\sum_{\substack{L \in \mathcal{D}'_m \\ K \times L \subset \Omega}} \sum_{\substack{I \in \mathcal{C}_K \\ J \in \mathcal{C}_L}} |\langle h_{I \times J}, T b_L^2 \rangle|^2 \lesssim \sum_{G \in \mathcal{G}_K} \mu(K \times G) \left(\frac{\ell(K)}{\ell(\tilde{K}_G)} \right)^{\alpha(1-\gamma_n)},$$

where \mathcal{G}_K consists of the maximal $G \in \mathcal{D}'_m$ such that $K \times G \subset \Omega$, and $\tilde{K}_G \in \mathcal{D}'_n$ is the maximal cube such that $\tilde{K}_G \supset K$ and $\tilde{K}_G \times G \subset \tilde{\Omega}$.

Before going to the proof, let us comment on the existence of the maximal cubes in the statement of the Lemma. Since Ω is bounded, there is an upper bound for $\ell(G)$ such that $K \times G \subset \Omega$, so \mathcal{G}_K is immediately well defined. As for \tilde{K}_G , there again arises the possibility of an infinite increasing sequence of $K_k \supset K$ with $K_k \times G \subset \tilde{\Omega}$. As before, such a sequence will exhaust \mathbb{R}^n , thus $\mathbb{R}^n \times G \subset \tilde{\Omega}$, and accordingly we interpret $\tilde{K}_G := \mathbb{R}^n$, $\ell(\tilde{K}_G) := \infty$ in this case. Note that the corresponding terms then vanish on the right of the asserted estimate.

Proof. Clearly every L as in the left is contained in a unique $G \in \mathcal{G}_K$; thus

$$\sum_{\substack{L \in \mathcal{D}'_m \\ K \times L \subset \Omega}} \sum_{\substack{I \in \mathcal{C}_K \\ J \in \mathcal{C}_L}} |\langle h_{I \times J}, T b_L^2 \rangle|^2 = \sum_{I \in \mathcal{C}_K} \sum_{G \in \mathcal{G}_K} \sum_{\substack{L \in \mathcal{D}'_m \\ L \subset G}} \sum_{J \in \mathcal{C}_L} |\langle h_{I \times J}, T b_L^2 \rangle|^2.$$

In this sum, we observe that $\tilde{K}_G \times L \subset \tilde{K}_G \times G \subset \tilde{\Omega}$, and hence by definition $\tilde{K}_G \subset F_L$. Now, for every $G \in \mathcal{G}_K$, there are two possibilities:

- There is at least one $L \subset G$ such that $F_L \neq \mathbb{R}^n$. Then $\tilde{K}_G \subset F_L$ is a proper dyadic cube.
- For all $L \subset G$, we have $F_L = \mathbb{R}^n$. But then for all these L , we are in the degenerate case with $b_L^2 \equiv 0$. Hence the part of the sum corresponding to these cubes G will vanish, and can be ignored.

Thus, we can restrict the summation to the cubes $G \in \mathcal{G}_K$ for which $\tilde{K}_G \in \mathcal{D}'_n$ is a proper dyadic cube.

Now, for summation variables as above, we have that $I \subset \hat{K} \subset K \subset \tilde{K}_G$ and $F_L^c \supset \tilde{K}_G^c$ are disjoint, and we may use the partial kernel representation

$$\begin{aligned} \langle h_{I \times J}, T b_L^2 \rangle &= \iint h_I(x_1) \langle h_J, T_1(x_1, y_1) b(y_1, \cdot) \rangle_{2\chi_{F_L^c}}(y_1) d\mu_n(x_1) d\mu_n(y_1) \\ &= \iint h_I(x_1) \langle h_J, B_I(x_1, y_1) \rangle_{2\chi_{F_L^c}}(y_1) d\mu_n(x_1) d\mu_n(y_1), \end{aligned}$$

where

$$B_I(x_1, y_1) := [T_1(x_1, y_1) - T_1(c_I, y_1)]b(y_1, \cdot).$$

We have

$$|c_I - y_1| \geq d(I, F_L^c) \geq d(I, K^c) \geq \ell(I)^{\gamma_n} \ell(K)^{1-\gamma_n} \geq \ell(I) \geq 2|x_1 - c_I|,$$

and therefore we may estimate

$$\|T_1(x_1, y_1) - T_1(c_I, y_1)\|_{CZ(\mu_m)} \lesssim \left(\frac{\ell(I)}{|c_I - y_1|} \right)^\alpha \frac{1}{\lambda_n(c_I, |c_I - y_1|)}.$$

Since $b(y_1, \cdot) \in L^\infty(\mu_m)$ is mapped into $\text{BMO}_3^2(\mu_m)$ by the Calderón–Zygmund operator above, we also have

$$\begin{aligned} \|B_I(x_1, y_1)\|_{\text{BMO}_3^2(\mu_m)} &\lesssim \|T_1(x_1, y_1) - T_1(c_I, y_1)\|_{CZ(\mu_m)} \\ &\lesssim \left(\frac{\ell(I)}{|c_I - y_1|}\right)^\alpha \frac{1}{\lambda_n(c_I, |c_I - y_1|)}. \end{aligned}$$

Estimating $\chi_{F_L^c}(y_1) \leq \chi_{\tilde{K}_G^c}(y_1)$, we then deduce that with a fixed $I \in \mathcal{C}_K$ there holds

$$\begin{aligned} &\sum_{G \in \mathcal{G}_K} \sum_{\substack{L \subset G \\ J \in \mathcal{C}_L}} |\langle h_{I \times J}, T b_L^2 \rangle|^2 \\ &\leq \left(\iint |h_I(x_1)| \left[\sum_{G \in \mathcal{G}_K} \chi_{\tilde{K}_G^c}(y_1) \sum_{\substack{L \subset G \\ J \in \mathcal{C}_L}} |\langle h_J, B_I(x_1, y_1) \rangle_2|^2 \right]^{1/2} d\mu_n(x_1) d\mu_n(y_1) \right)^2. \end{aligned}$$

We then note that

$$\begin{aligned} \sum_{\substack{L \subset G \\ J \in \mathcal{C}_L}} |\langle h_J, B_I(x_1, y_1) \rangle_2|^2 &\leq \sum_{\substack{J \text{ good} \\ J \subset G \\ \ell(J) \leq 2^{-r}\ell(G)}} |\langle h_J, B_I(x_1, y_1) \rangle_2|^2 \\ &\lesssim \mu_m(G) \|B_I(x_1, y_1)\|_{\text{BMO}_3^2(\mu_m)}^2 \end{aligned}$$

by Lemma 2.10. This gives us that

$$\begin{aligned} &\sum_{G \in \mathcal{G}_K} \sum_{\substack{L \subset G \\ J \in \mathcal{C}_L}} |\langle h_{I \times J}, T b_L^2 \rangle|^2 \\ &\lesssim \left(\iint |h_I(x_1)| \|B_I(x_1, y_1)\|_{\text{BMO}_3^2(\mu_m)} \left[\sum_{G \in \mathcal{G}_K} \mu_m(G) \chi_{\tilde{K}_G^c}(y_1) \right]^{1/2} d\mu_n(x_1) d\mu_n(y_1) \right)^2 \\ &\lesssim \mu_n(I) \left(\int \left(\frac{\ell(I)}{|c_I - y_1|} \right)^\alpha \frac{1}{\lambda_n(c_I, |c_I - y_1|)} \left[\sum_{G \in \mathcal{G}_K} \mu_m(G) \chi_{\tilde{K}_G^c}(y_1) \right]^{1/2} d\mu_n(y_1) \right)^2. \end{aligned}$$

The integral squared is estimated by

$$\begin{aligned} &\left(\int_{I^c} \left(\frac{\ell(I)}{|c_I - y_1|} \right)^\alpha \frac{d\mu_n(y_1)}{\lambda_n(c_I, |c_I - y_1|)} \right) \\ &\quad \times \left(\int \left(\frac{\ell(I)}{|c_I - y_1|} \right)^\alpha \frac{1}{\lambda_n(c_I, |c_I - y_1|)} \sum_{G \in \mathcal{G}_K} \mu_m(G) \chi_{\tilde{K}_G^c}(y_1) d\mu_n(y_1) \right), \end{aligned}$$

which is further dominated by

$$\begin{aligned}
& \sum_{G \in \mathcal{G}_K} \mu_m(G) \int_{\tilde{K}_G^c} \left(\frac{\ell(I)}{|c_I - y_1|} \right)^\alpha \frac{d\mu_n(y_1)}{\lambda_n(c_I, |c_I - y_1|)} \\
& \lesssim \sum_{G \in \mathcal{G}_K} \mu_m(G) \int_{y_1: d(\hat{K}, y_1) \geq 2^{-r\gamma_n} \ell(K)^{\gamma_n} \ell(\tilde{K}_G)^{1-\gamma_n}} \left(\frac{\ell(K)}{d(\hat{K}, y_1)} \right)^\alpha \frac{d\mu_n(y_1)}{\lambda_n(x_{\hat{K}}, d(\hat{K}, y_1))} \\
& \lesssim \sum_{G \in \mathcal{G}_K} \mu_m(G) \left(\frac{\ell(K)}{\ell(\tilde{K}_G)} \right)^{\alpha(1-\gamma_n)}.
\end{aligned}$$

These estimates follow completely analogously to the ones in the previous subsection and utilize goodness in an essential way.

We have established the bound

$$\begin{aligned}
& \sum_{I \in \mathcal{C}_K} \sum_{G \in \mathcal{G}_K} \sum_{\substack{L \subset G \\ J \in \mathcal{C}_L}} |\langle h_{I \times J}, T b_L^2 \rangle|^2 \lesssim \sum_{I \in \mathcal{C}_K} \mu_n(I) \sum_{G \in \mathcal{G}_K} \mu_m(G) \left(\frac{\ell(K)}{\ell(\tilde{K}_G)} \right)^{\alpha(1-\gamma_n)} \\
& = \mu_n(\hat{K}) \sum_{G \in \mathcal{G}_K} \mu_m(G) \left(\frac{\ell(K)}{\ell(\tilde{K}_G)} \right)^{\alpha(1-\gamma_n)} \\
& \leq \sum_{G \in \mathcal{G}_K} \mu(K \times G) \left(\frac{\ell(K)}{\ell(\tilde{K}_G)} \right)^{\alpha(1-\gamma_n)},
\end{aligned}$$

which ends the proof of the lemma. \square

It remains to estimate the right side of the Lemma summed over $K \in \mathcal{D}'_n$:

$$\begin{aligned}
& \sum_{K \in \mathcal{D}'_n} \sum_{G \in \mathcal{G}_K} \mu(K \times G) \left(\frac{\ell(K)}{\ell(\tilde{K}_G)} \right)^{\alpha(1-\gamma_n)} \leq \sum_{\substack{K \times G \subset \Omega \\ \text{2-maximal}}} 2^{-\alpha(1-\gamma_n) \text{emb}_1(K \times G; \Omega)} \mu(K \times G) \\
& \leq 2 \sum_{k=0}^{\infty} 2^{-\alpha(1-\gamma_n)k} \times \mu(\Omega) \lesssim \mu(\Omega).
\end{aligned}$$

Here we used that by the definition of \mathcal{G}_K , the rectangles $K \times G \subset \Omega$ are 2-maximal. Second, by the definition of \tilde{K}_G we have

$$\frac{\ell(K)}{\ell(\tilde{K}_G)} = 2^{-\text{emb}_1(K \times G; \Omega)}.$$

Lastly, we utilized Theorem 8.1. This completes the proof of Theorem 8.5.

REFERENCES

- [1] S. Y. A. Chang, R. Fefferman, A continuous version of duality of H^1 with BMO on the Bidisc, Ann. of Math. 112 (1) (1980) 179–201.
- [2] S. Y. A. Chang, R. Fefferman, Some recent developments in Fourier analysis and H^p theory on product domains, Bull. Amer. Math. Soc. 12 (1) (1985) 1–43.

- [3] C. Cabrelli, M. Lacey, U. Molter, J. Pipher, Variations on the theme of Journé's lemma, *Houston J. Math.* 32 (3) (2006) 833–861.
- [4] G. David, J.-L. Journé, A boundedness criterion for generalized Calderón-Zygmund operators, *Ann. of Math.* 120 (2) (1984) 371–397.
- [5] G. David, J.-L. Journé, S. Semmes, Opérateurs de Calderón-Zygmund, fonctions para-accretives et interpolation, *Rev. Mat. Iberoamericana* 1 (4) (1985) 1–56.
- [6] R. Fefferman, Harmonic analysis on product spaces, *Ann. of Math.* 126 (1) (1987) 109–130.
- [7] R. Fefferman, E. Stein, Singular integrals on product spaces, *Adv. Math.* 45 (2) (1982) 117–143.
- [8] T. Hytönen, Representation of singular integrals by dyadic operators, and the A_2 theorem, *arXiv:1108.5119*, 2011.
- [9] T. Hytönen, The sharp weighted bound for general Calderón-Zygmund operators, *Ann. of Math.* 175 (3) (2012) 1473–1506.
- [10] T. Hytönen, The vector-valued non-homogeneous Tb theorem, *Int. Math. Res. Not. IMRN* (2012), in press, *arXiv:0809.3097*, 2009.
- [11] T. Hytönen, H. Martikainen, Non-homogeneous Tb theorem and random dyadic cubes on metric measure spaces, *J. Geom. Anal.* 22 (4) (2012) 1071–1107.
- [12] T. Hytönen, D. Yang, D. Yang, The Hardy space H^1 on non-homogeneous metric spaces, *Math. Proc. Cambridge Philos. Soc.* 153 (1) (2012) 9–31.
- [13] J.-L. Journé, A covering lemma for product spaces, *Proc. Amer. Math. Soc.* 96 (4) (1986) 593–598.
- [14] J.-L. Journé, Calderón-Zygmund operators on product spaces, *Rev. Mat. Iberoamericana* 1 (3) (1985) 55–91.
- [15] M. Lacey, J. Metcalfe, Paraproducts in one and several parameters, *Forum Math.* 19 (2) (2007) 325–351.
- [16] H. Martikainen, Representation of bi-parameter singular integrals by dyadic operators, *Adv. Math.* 229 (3) (2012) 1734–1761.
- [17] H. Martikainen, Vector-valued non-homogeneous Tb theorem on metric measure spaces, *Rev. Mat. Iberoamericana* (2012), in press, *arXiv:1004.3176*, 2010.
- [18] F. Nazarov, S. Treil, A. Volberg, Accretive system Tb -theorems on nonhomogeneous spaces, *Duke Math. J.* 113 (2) (2002) 259–312.
- [19] F. Nazarov, S. Treil, A. Volberg, The Tb -theorem on non-homogeneous spaces, *Acta Math.* 190 (2) (2003) 151–239.
- [20] S. Pott, P. Villarroya, A $T(1)$ theorem on product spaces, preprint, *arXiv:1105.2516*, 2011.
- [21] J. Pipher, L. A. Ward, BMO from dyadic BMO on the bidisc, *J. London Math. Soc.* 77 (2) (2008) 524–544.
- [22] S. Treil, H^1 and dyadic H^1 , in: Linear and complex analysis, in: *Amer. Math. Soc. Transl. Ser. 2*, vol. 226, Amer. Math. Soc., Providence, RI, 2009, pp. 179–193.

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF HELSINKI, P.O.B. 68, FI-00014 HELSINKI, FINLAND

E-mail address: tuomas.hytonen@helsinki.fi

DÉPARTEMENT DE MATHÉMATIQUES, BÂTIMENT 425, FACULTÉ DES SCIENCES D'ORSAY, UNIVERSITÉ PARIS-SUD 11, F-91405 ORSAY CEDEX

E-mail address: henri.martikainen@math.u-psud.fr